



```
QY 2 PKPOQWFW 9
   |||:|
Db 328 PKPSEWVW 335

RESULT 12
ATP8_SQUAC
ID ATP8_SQUAC STANDARD; PRT; 55 AA.
AC Q92250;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).
GN MTATP8 OR ATP8.
OS Squalus acanthias (Spiny dogfish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99091711; PubMed=9873084;
RA Rasmussen A.S., Arnason U.;
RT "Phylogenetic studies of complete mitochondrial DNA molecules place
RT cartilaginous fishes within the tree of bony fishes.";
RL J. Mol. Evol. 48:118-123(1999).
CC -1- FUNCTION: THIS IS ONE OF THE CHAINS OF THE NONENZYMATIC COMPONENT
CC (CF(0) SUBUNIT) OF THE MITOCHONDRIAL ATPASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.
CC -1- SIMILARITY: BELONGS TO THE ATPASE PROTEIN 8 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; Y18134; CAA77053.1; -
DR InterPro; IPR001421; ATP-synt_8.
DR Pfam; PF00895; ATP-synt_8; 1.
DR K W Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT TRANSMEM 6 26 POTENTIAL.
SQ SEQUENCE 55 AA; 6587 MW; 3FB9F843CEFA54EE CRC64;

Query Match 57.7%; Score 41; DB 1; Length 55;
Best Local Similarity 55.6%; Pred. No. 6.8;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPOQWFW 9
   |||:|
Db 44 KPKPEPWNW 52

RESULT 13
CIW2_MOUSE
ID CIW2_MOUSE STANDARD; PRT; 411 AA.
AC P97438;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE POTASSIUM CHANNEL SUBFAMILY K MEMBER 2 (OUTWARD RECTIFYING POTASSIUM
DE CHANNEL PROTEIN TREK-1) (TWO-PORE POTASSIUM CHANNEL TPCK1) (TREK-1 K+
DE CHANNEL SUBUNIT).
GN KCNK2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.
RP TISSUE=Brain;
RX MEDLINE=97157476; PubMed=9003761;
RA Fink M., Duprat F., Lesage F., Reyes R., Romey G., Heurteaux C.,
RA Lazdunski M.;
RT "Cloning, functional expression and brain localization of a novel
RT unconventional outward rectifier K+ channel.";
RL EMBO J. 15:6854-6862(1996).
RN [2]
RP REVISIONS
RP TISSUE=Brain;
RA Fink M., Duprat F., Lesage F., Reyes R., Romey G., Heurteaux C.,
RA Lazdunski M.;
RN Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP ACTIVATION
RX MEDLINE=99254548; PubMed=10321245;
RA Patel A.J., Honore E., Lesage F., Fink M., Romey G., Lazdunski M.;
RT "Inhalational anesthetics activate two-pore-domain background K+
RT channels.";
RL Nat. Neurosci. 2:422-426(1999).
CC -1- FUNCTION: OUTWARD RECTIFYING POTASSIUM CHANNEL.
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN BRAIN AND LUNG. ALSO
CC DETECTED IN KIDNEY, HEART AND SKELETAL MUSCLE. NOT DETECTED IN
CC LIVER. IN THE BRAIN, HIGHEST EXPRESSION IN OLFACTORY BULB,
CC HIPPOCAMPUS AND CEREBELLUM.
CC -1- MISCELLANEOUS: INHIBITED BY BARIUM. ACTIVATED BY VOLATILE GENERAL
CC ANAESTHETICS SUCH AS CHLOROFORM, DIETHYL ETHER, HALOTHANE AND
CC ISOFLURANE.
CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM
CC CHANNELS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U73488; AAC53005.2; -
DR MGD; MGI:109366; Kcnk2.
DR InterPro; IPR003280; 2poreK_channel.
DR InterPro; IPR001622; Channel_pore_K.
DR InterPro; IPR000099; TWIK_channel.
DR Pfam; PF02034; TWIK_channel; 1.
DR PRINTS; PR01333; 2PORECHANNEL.
DR Ionic channel; Transmembrane; Ion transport; Potassium transport;
KW Glycoprotein.
FT DOMAIN 1 46 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 47 67 POTENTIAL.
FT DOMAIN 129 155 PORE-FORMING (POTENTIAL).
FT TRANSMEM 157 177 POTENTIAL.
FT DOMAIN 178 207 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 208 228 POTENTIAL.
FT DOMAIN 238 268 PORE-FORMING (POTENTIAL).
FT TRANSMEM 273 293 POTENTIAL.
FT DOMAIN 294 411 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 378 411 ESSENTIAL FOR CHLOROFORM AND HALOTHANE
SENSITIVITY.
FT DOMAIN 354 411 REQUIRED FOR BASAL CHANNEL ACTIVITY.
FT CARBOHYD 95 95 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 119 119 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 411 AA; 45297 MW; 8F976DD103EFA05 CRC64;

Query Match 56.3%; Score 40; DB 1; Length 411;
Best Local Similarity 55.6%; Pred. No. 58;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQOWFWLM 11
```



```
FT CARBOHYD 148 148 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 538 AA; 59764 MW; 8EA615B08D147FBC CRC64;

Query Match 56.3%; Score 40; DB 1; Length 538;
Best Local Similarity 55.6%; Pred. No. 74;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWM 11
   || |||::
DB 297 KPLVWFIL 305

RESULT 16
CIWA_RAT STANDARD; PRT; 538 AA.
AC Q9JIS4;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE POTASSIUM CHANNEL SUBFAMILY K MEMBER 10 (OUTWARD RECTIFYING POTASSIUM CHANNEL PROTEIN TREK-2) (TREK-2 K+ CHANNEL SUBUNIT).
GN KCNK10 OR TREK2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20298807; PubMed=10747911;
RA Bang H., Kim Y., Kim D.;
RT "TREK-2, a new member of the mechanosensitive tandem-pore K+ channel family."
RL J. Biol. Chem. 275:17412-17419(2000).
CC -1- FUNCTION: OUTWARD RECTIFYING POTASSIUM CHANNEL. PRODUCES RAPIDLY ACTIVATING AND NON-INACTIVATING OUTWARD RECTIFIER K(+) CURRENTS.
CC ACTIVATED BY ARACHIDONIC ACID AND OTHER NATURALLY OCCURRING UNSATURATED FREE FATTY ACIDS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN THE CEREBELLUM, SPLEEN, AND TESTIS.
CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM CHANNELS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
-----
EMBL; AF196965; AAF75132.1; -
DR InterPro; IPR003280; 2poreK_channel.
DR InterPro; IPR001622; Channel_pore_K.
DR InterPro; IPR000099; TWIK_channel.
DR Pfam; PF02034; TWIK_channel; 1.
DR PRINTS; PR01333; 2PORECHANNEL.
KW Ionic channel; Transmembrane; Ion transport; Potassium transport; Glycoprotein.
FT DOMAIN 1 71 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 72 92 POTENTIAL.
FT DOMAIN 154 180 PORE-FORMING (BY SIMILARITY).
FT TRANSMEM 182 202 POTENTIAL.
FT DOMAIN 203 233 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 234 254 POTENTIAL.
FT DOMAIN 263 294 PORE-FORMING (BY SIMILARITY).
FT TRANSMEM 299 319 POTENTIAL.
FT DOMAIN 320 538 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 144 144 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 147 147 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 538 AA; 59800 MW; 1FF33F0AA52B97E4 CRC64;
```

```
Query Match 56.3%; Score 40; DB 1; Length 538;
Best Local Similarity 55.6%; Pred. No. 74;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWM 11
   || |||::
DB 297 KPLVWFIL 305

RESULT 17
CNG_DROME STANDARD; PRT; 665 AA.
AC Q24278;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE CYCLIC-NUCLEOTIDE-GATED CATION CHANNEL (CNG CHANNEL).
GN CNG.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95045396; PubMed=7957070;
RA Baumann A., Frings S., Godde M., Seifert R., Kaupp U.B.;
RT "Primary structure and functional expression of a Drosophila cyclic nucleotide-gated channel present in eyes and antennae."
RL EMBO J. 13:5040-5050(1994).
CC -1- FUNCTION: APPROXIMATELY 50-FOLD MORE SENSITIVE TO cGMP THAN TO CAMP. MAY BE INVOLVED IN TRANSDUCTION CASCADES OF BOTH INVERTEBRATE PHOTORECEPTORS AND OLFACTORY SENSILLAE.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ANTENNAE AND THE VISUAL SYSTEM.
CC -1- SIMILARITY: BELONGS TO THE CYCLIC NUCLEOTIDE-GATED CATION CHANNEL FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch).
-----
EMBL; X89601; CAA61760.1; -
DR FlyBase; FBgn0014462; Cng.
DR InterPro; IPR002025; CNG_membrane.
DR InterPro; IPR000595; cNMP_binding.
DR Pfam; PF00914; CNG_membrane; 1.
DR Pfam; PF00027; cNMP_binding; 1.
DR SMART; SM00100; cNMP; 1.
DR PROSITE; PS00888; cNMP_BINDING_1; 1.
DR PROSITE; PS00889; cNMP_BINDING_2; 1.
DR PROSITE; PS00042; cNMP_BINDING_3; 1.
DR KW Ionic channel; Ion transport; cAMP-binding; Transmembrane.
FT DOMAIN 1 110 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 111 131 H1 (POTENTIAL).
FT DOMAIN 132 139 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 139 159 H2 (POTENTIAL).
FT DOMAIN 160 186 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 187 207 H3 (POTENTIAL).
FT DOMAIN 208 253 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 254 274 H4 (POTENTIAL).
FT DOMAIN 275 325 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 326 346 H5 (POTENTIAL).
FT DOMAIN 347 481 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 482 502 H6 (POTENTIAL).
FT DOMAIN 503 665 CYTOPLASMIC (POTENTIAL).
FT NP_BIND 437 559 CAMP (BY SIMILARITY).
FT BINDING 496 496 CAMP (POTENTIAL).
FT BINDING 511 511 CAMP (POTENTIAL).
```



FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 665 AA; 75922 MW; 9FLBDC5D9581C8DB CRC64;

Query Match 56.3%; Score 40; DB 1; Length 665;  
Best Local Similarity 75.0%; Pred. NO. 90;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 RPKPOQWF 8  
||||| II  
Db 52 RKPDPWF 59

RESULT 18  
FUT3 HUMAN  
ID FUT3\_HUMAN STANDARD; PRT; 361 AA.  
AC P21217; Q99448; Q99449;  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE GALACTOSIDE 3(4)-L-FUCOSYLTRANSFERASE (EC 2.4.1.65) (BLOOD GROUP LEWIS  
DE ALPHA-4-FUCOSYLTRANSFERASE) (LEWIS FT) (FUCOSYLTRANSFERASE 3) (FUC-  
DE I11).  
GN FUT3 OR LE.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91032981; PubMed=1977660;  
RA Kulkowski-Latallo J.F., Larsen R.D., Nair R.P., Lowe J.B.;  
RT "A cloned human cDNA determines expression of a mouse stage-specific  
RT embryonic antigen and the Lewis blood group  
RT alpha(1,3/1,4)fucosyltransferase.";  
RL Genes Dev. 4:1288-1303(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=95378269; PubMed=7650030;  
RA Cameron H.S., Szczepaniak D., Weston W.;  
RT "Expression of human chromosome 19p alpha(1,3)-fucosyltransferase  
RT genes in normal tissues. Alternative splicing, polyadenylation, and  
RT isoforms.";  
RL J. Biol. Chem. 270:20112-20122(1995).  
RN [3]  
RP VARIANT LE(-) MET-105.  
RX MEDLINE=94059067; PubMed=8240322;  
RA Elmgren A., Rydberg L., Larson G.;  
RT "Genotypic heterogeneity among Lewis negative individuals.";  
RL Biochem. Biophys. Res. Commun. 196:515-520(1993).  
RN [4]  
RP VARIANT LE(-) ARG-20; SER-170 AND ALA-336.  
RX MEDLINE=94059082; PubMed=8240337;  
RA Nishihara S., Yazawa S., Iwasaki H., Nakazato M., Kudo T., Ando T.,  
RA Narimatsu H.;  
RT "Alpha (1,3/1,4)fucosyltransferase (FucT-III) gene is inactivated by  
RT a single amino acid substitution in Lewis histo-blood type negative  
RT individuals.";  
RL Biochem. Biophys. Res. Commun. 196:624-631(1993).  
RN [5]  
RP VARIANT LE(-) ARG-20 AND SER-170.  
RX MEDLINE=94033579; PubMed=8219240;  
RA Koda Y., Kimura H., Mekada E.;  
RT "Analysis of Lewis fucosyltransferase genes from the human gastric  
RT mucosa of Lewis-positive and -negative individuals.";  
RL Blood 82:2915-2919(1993).  
RN [6]  
RP VARIANT LE(-) ARG-20 AND LYS-356.  
RX MEDLINE=94342259; PubMed=8063716;  
RA Mollicone R., Reuguine I., Kelly R.J., Fletcher A., Watt J.,  
RA Chatfield S., Aziz A., Cameron H.S., Weston B.W., Lowe J.B., Oriol R.;  
RT "Molecular basis for Lewis alpha(1,3/1,4)-fucosyltransferase gene

deficiency (FUT3) found in Lewis-negative Indonesian pedigrees.";  
J. Biol. Chem. 269:20987-20994(1994).  
RN [7]  
RP VARIANT LE(-) LYS-356.  
RX MEDLINE=95050753; PubMed=7961897;  
RA Nishihara S., Narimatsu H., Iwasaki H., Yazawa S., Akamatsu S.,  
RA Ando T., Seno T., Narimatsu I.;  
RT "Molecular genetic analysis of the human Lewis histo-blood group  
RT system.";  
RL J. Biol. Chem. 269:29271-29278(1994).  
RN [8]  
RP VARIANT LE(-) ARG-20; ARG-68; MET-105 AND LYS-356.  
RX MEDLINE=96243526; PubMed=8801770;  
RA Elmgren A., Boerjeson C., Svensson L., Rydberg L., Larson G.;  
RT "DNA sequencing and screening for point mutations in the human Lewis  
RT 'FUT3' gene enables molecular genotyping of the human Lewis blood  
RT group system.";  
RL Vox Sang. 70:97-103(1996).  
RN [9]  
RP VARIANT LE(-) ARG-68 AND MET-105.  
RX MEDLINE=97413801; PubMed=9268337;  
RA Elmgren A., Mollicone R., Costache M., Boerjeson C., Oriol R.,  
RA Harrington J., Larson G.;  
RT "Significance of individual point mutations, T202C and C314T, in the  
RT human Lewis 'FUT3' gene for expression of Lewis antigens by the human  
RT alpha'1,3/1,4'-fucosyltransferase, Fuc-TIII.";  
RL J. Biol. Chem. 272:21994-21998(1997).  
RN [10]  
RP VARIANT LE(+) K-102; A-124, AND VARIANTS LE(-) N-162; R-223; M-270.  
RX MEDLINE=98366989; PubMed=9703429;  
RA Pang H., Liu Y., Koda Y., Soejima M., Jia J., Schlaphoff T.,  
RA du Toit E.D., Kimura H.;  
RT "Five novel missense mutations of the Lewis gene 'FUT3' in African  
RT 'Xhosa' and Caucasian populations in South Africa.";  
RL Hum. Genet. 102:675-680(1998).  
CC -!- FUNCTION: MAY CATALYZE ALPHA-1,3 AND ALPHA-1,4 GLYCOSIDIC LINKAGES  
CC INVOLVED IN THE EXPRESSION OF VIM-2, LEWIS A, LEWIS B, STALYL  
CC LEWIS X AND LEWIS X/SSEA-1 ANTIGENS, MAY BE INVOLVED IN BLOOD  
CC GROUP LEWIS DETERMINATION; LEWIS-POSITIVE (LE(+)) INDIVIDUALS  
CC HAVE AN ACTIVE ENZYME WHILE LEWIS-NEGATIVE (LE(-)) INDIVIDUALS  
CC HAVE AN INACTIVE ENZYME.  
CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + 1,3-BETA-D-GALACTOSYL-  
CC N-ACETYL-D-GLUCOSAMINYL-R = GDP + 1,3-BETA-D-GALACTOSYL-  
CC (ALPHA-1,4-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINYL-R.  
CC -!- PATHWAY: GLYCOSYLATION.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
CC FORM IN TRANS CISTERNAE OF GOLGI.  
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN STOMACH, COLON, SMALL  
CC INTESTINE, LUNG AND KIDNEY AND TO A LESSER EXTENT IN SALIVARY  
CC GLAND, BLADDER, UTERUS AND LIVER.  
CC -!- MISCELLANEOUS: ALSO ACTS ON THE CORRESPONDING 1,4-GALACTOSYL  
CC DERIVATIVE, FORMING 1,3-L-FUCOSYL LINKS.  
CC -!- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN  
CC GLYCOSYLTRANSFERASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X53578; CAA37641.1; -  
DR EMBL; U27328; AAC50187.1; -  
DR EMBL; U27326; AAC50185.1; -  
DR EMBL; U27327; AAC50186.1; -  
DR EMBL; D89324; BAAL3941.1; -  
DR EMBL; D89325; BAAL3942.1; -  
DR EMBL; A36669; A36669.  
DR MIR; 111100; -  
DR InterPro; IPR001503; Glyco\_transf\_10.  
DR Pfam; PF00852; Glyco\_transf\_10; 1.

KW Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;  
KW Signal-anchor; Golgi stack; Polymorphism; Blood group antigen.  
FT DOMAIN 1 15  
FT TRANSMEM 16 34  
FT TRANSMEM 35 361  
FT DOMAIN 154 154  
FT CARBOHYD 185 185  
FT CARBOHYD 20 20  
FT VARIANT 68 68  
FT VARIANT 102 102  
FT VARIANT 105 105  
FT VARIANT 124 124  
FT VARIANT 162 162  
FT VARIANT 170 170  
FT VARIANT 223 223  
FT VARIANT 270 270  
FT VARIANT 336 336  
FT VARIANT 356 356  
FT SEQUENCE 361 AA; 42117 MW; BF4398044F19C284 CRC64;

Query Match 54.98; Score 39; DB 1; Length 361;  
Best Local Similarity 55.68; Pred. No. 71;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 RPKPQQWF 9  
Db 126 RPOGQRIW 134

RESULT 19  
FUT3\_PANTR STANDARD; PRT; 372 AA.  
AC O19058;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE GALACTOSIDE 3(4)-L-FUCOSYLTRANSFERASE (EC 2.4.1.65) (BLOOD GROUP LEWIS  
DE ALPHA-4-FUCOSYLTRANSFERASE) (LEWIS PT) (FUCOSYLTRANSFERASE 3) (FUCT-  
DE III) (ALPHA-3/4-FUCOSYLTRANSFERASE).  
GN FUT3.  
OS Pan troglodytes (Chimpanzee).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
OX NCBI\_TaxID=9598;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98037800; PubMed=9368041;  
RA Costache M., Apoll P.-A., Calileau A., Elmgren A., Larson G.,  
RA Henry S., Blancher A., Iordachescu D., Oriol R., Mollicone R.;  
RA "Evolution of fucosyltransferase genes in vertebrates.";  
RL J. Biol. Chem. 272:29721-29728(1997).  
CC -1- FUNCTION: MAY CATALYZE ALPHA-1,3 AND ALPHA-1,4 GLYCOSIDIC LINKAGES  
CC INVOLVED IN THE EXPRESSION OF STIALYL LEWIS X AND LEWIS X/SSEA-1  
CC ANTIGENS. IT MAY BE INVOLVED IN BLOOD GROUP LEWIS DETERMINATION  
CC (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: GDP-L-FUCOSE + 1,3-BETA-D-GALACTOSYL-  
CC N-ACETYL-D-GLUCOSAMINY-L-R = GDP + 1,3-BETA-D-GALACTOSYL-  
CC (ALPHA-1,4-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINY-L-R.  
CC -1- PATHWAY: GLYCOSYLATION.  
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND

CC FORM IN TRANS CISTERNAE OF GOLGI (BY SIMILARITY).  
CC -1- POLYMORPHISM: THERE ARE TWO ALLELES (A AND B). ALLELE A HAS ARG-  
CC 162 AND VAL-304. ALLELE B HAS GLY-162 AND MET-304.  
CC -1- MISCELLANEOUS: ALSO ACTS ON THE CORRESPONDING 1,4-GALACTOSYL  
CC DERIVATIVE, FORMING 1,3-L-FUCOSYL LINKS.  
CC -1- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN  
CC GLYCOSYLTRANSFERASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Y14033; CAA74360.1;  
DR InterPro: IPR001503; Glyco\_transf\_10.  
DR Pfam: PF00852; Glyco\_transf\_10; 1.  
KW Transferrase; Glycosyltransferase; Glycoprotein; Transmembrane;  
KW Signal-anchor; Golgi stack; Polymorphism.  
FT DOMAIN 1 14  
FT TRANSMEM 15 34  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT DOMAIN 35 372  
FT LUMENAL, CATALYTIC (POTENTIAL).  
FT CARBOHYD 165 165  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 196 196  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 162 162  
FT R -> G (IN ALLELE B).  
FT VARIANT 304 304  
FT V -> M (IN ALLELE B).  
FT SEQUENCE 372 AA; 43233 MW; 649CBFBBCA7BD74C CRC64;

Query Match 54.98; Score 39; DB 1; Length 372;  
Best Local Similarity 55.68; Pred. No. 73;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 RPKPQQWF 9  
Db 137 RPOGQRIW 145

RESULT 20  
FUT5\_HUMAN STANDARD; PRT; 374 AA.  
AC Q11128;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE ALPHA-(1,3)-FUCOSYLTRANSFERASE (EC 2.4.1.65) (GALACTOSIDE 3-L-  
DE FUCOSYLTRANSFERASE) (FUCOSYLTRANSFERASE 5) (FUCT-V).  
GN FUT5.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Peripheral blood leukocytes;  
RX MEDLINE=92156161; PubMed=1740457;  
RA Weston B.W., Nair R.P., Larsen R.D., Lowe J.B.;  
RA "Isolation of a novel human alpha (1,3)fucosyltransferase gene and  
RA molecular comparison to the human Lewis blood group alpha  
RA (1,3/4)fucosyltransferase gene. Syntenic, homologous, nonallelic  
RA genes encoding enzymes with distinct acceptor substrate  
RA specificities.";  
RL J. Biol. Chem. 267:4152-4160(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Colon, Kidney, and Liver;  
RX MEDLINE=95378269; PubMed=7650030;  
RA Cameron H.S., Szczepaniak D., Weston B.W.;  
RA "Expression of human chromosome 19p alpha(1,3)-fucosyltransferase  
RA genes in normal tissues. Alternative splicing, polyadenylation, and



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; M17522; AAA25572.1; -;  
DR EMBL; X05799; CAA29244.1; -;  
DR PIR; B29413; B29413.  
DR InterPro; IPR000179; Cyt\_b6.  
DR Pfam; PF00032; cytochrome\_b\_6; 2.  
DR Pfam; PF00033; cytochrome\_b\_6; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
DR PROSITE; PS00193; CYTOCHROME\_B\_OO; 1.  
KW Electron transport; Respiratory chain; Heme; Transmembrane.  
FT METAL 97 97  
FT METAL 111 111 IRON 1 (HEME B562 AXIAL LIGAND).  
FT METAL 198 198 IRON 2 (HEME B566 AXIAL LIGAND).  
FT METAL 212 212 IRON 1 (HEME B566 AXIAL LIGAND).  
SQ SEQUENCE 440 AA; 50116 MW; 8D211B8614920C63 CRC64;

Query Match 54.9%; Score 39; DB 1; Length 440;  
Best Local Similarity 54.5%; Pred. No. 85;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKQOQFWLM 11  
|| : ||||:  
Db 360 RPLFKWFWLL 370

## RESULT 23

NPAA2\_MOUSE STANDARD; PRT; 816 AA.  
AC P97460;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE NEURONAL PAS DOMAIN PROTEIN 2 (NEURONAL PAS2).  
GN NPAS2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=97165088; PubMed=9012850;  
RA Zhou Y.-D., Barnard M., Tian H., Li X., Ring H.Z., Francke U.,  
RA Shelton J., Richardson J., Russell D.W., McKnight S.L.;  
RT "Molecular characterization of two mammalian bHLH-PAS domain proteins  
selectively expressed in the central nervous system.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:713-718(1997).  
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER  
CC bHLH PROTEIN. INTERACTS WITH HSP90.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
CC -1- TISSUE SPECIFICITY: IN BRAIN, EXCLUSIVELY NEURONAL. ALSO FOUND IN  
CC SPINAL CORD, AND IN A LESSER EXTENT IN COLON, SMALL INTESTINE AND  
CC UTERUS.  
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED 3 DAYS AFTER BIRTH.  
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
CC TRANSCRIPTION FACTORS.  
CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U77969; AAB47249.1; -;  
DR MGD; MGI:109232; Npas2.  
DR InterPro; IPR003015; HLH\_Myc.  
DR InterPro; IPR001092; HLH\_dim.  
DR InterPro; IPR001610; PAC.  
DR InterPro; IPR000014; PAC.  
DR Pfam; PF00010; HLH; 1.  
DR Pfam; PF00785; PAC; 1.  
DR Pfam; PF00989; PAC; 2.  
DR SMART; SM00353; HLH; 1.  
DR SMART; SM00086; PAC; 1.  
DR SMART; SM00091; PAS; 2.  
DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; 1.  
KW Repeat; DNA-binding; Nuclear protein; Transcription regulation.  
FT DNA\_BIND 10 22 BASIC DOMAIN.  
FT DOMAIN 23 60 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).  
FT REPEAT 84 150 PAS-1.  
FT REPEAT 239 305 PAS-2.  
FT DOMAIN 311 354 PAC MOTIF.  
SQ SEQUENCE 816 AA; 90915 MW; 7E5CF0641CFDC1DD CRC64;

Query Match 54.9%; Score 39; DB 1; Length 816;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQFWL 10  
| ||| ||  
Db 319 KGQQWIL 326

## RESULT 24

NPAA2\_HUMAN STANDARD; PRT; 824 AA.  
AC Q99743; Q99629;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE NEURONAL PAS DOMAIN PROTEIN 2 (NEURONAL PAS2) (MEMBER OF PAS PROTEIN  
DE 4) (MOP4).  
GN NPAS2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97165088; PubMed=9012850;  
RA Zhou Y.-D., Barnard M., Tian H., Li X., Ring H.Z., Francke U.,  
RA Shelton J., Richardson J., Russell D.W., McKnight S.L.;  
RT "Molecular characterization of two mammalian bHLH-PAS domain proteins  
selectively expressed in the central nervous system.";  
RL Proc. Natl. Acad. Sci. U.S.A. 94:713-718(1997).  
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER  
CC bHLH PROTEIN. INTERACTS WITH HSP90.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
CC TRANSCRIPTION FACTORS.  
CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U77970; AAB47250.1; -.
DR EMBL; U51625; AAC51211.1; -.
DR MIM; 603347; -.
DR InterPro; IPR003015; HLH_MYC.
DR InterPro; IPR001092; HLH_dlm.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
DR Repeat; DNA-binding; Nuclear protein; Transcription regulation.
KW DNA_BIND 10 22
FT DOMAIN 23 60
FT REPEAT 84 150
FT REPEAT 239 305
FT DOMAIN 311 354
FT CONFLICT 51 51
FT CONFLICT 164 164
FT CONFLICT 308 308
FT CONFLICT 471 471
SQ SEQUENCE 824 AA; 91759 MW; 249A4C687B328A5 CRC64;

Query Match 54.9%; Score 39; DB 1; Length 824;
Best Local Similarity 75.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQWFWL 10
DB 319 KGQWFWL 326

RESULT 25
CLOC_HUMAN
ID CLOC_HUMAN STANDARD; PRT; 846 AA.
AC O15516; O14516; Q9UIT8;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CIRCADIAN LOCOMOTOR OUTPUT CYCLES KAPUT PROTEIN (HCLOCK).
GN CLOC OR KIA0334.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA King D.P., Steeves T.D.L., Zhao Y., Sangoram A.M., Takahashi J.S.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBSJ databases.
RW [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99216412; PubMed=10198158;
RA Steeves T.D.L., King D.P., Zhao Y., Sangoram A.M., Du F.,
RA Bowcock A.M., Moore R.Y., Takahashi J.S.;
RT "Molecular cloning and characterization of the human CLOCK gene:
RT expression in the suprachiasmatic nuclei.";
RL Genomics 57:189-200(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
```

```
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
RN [4]
RP SEQUENCE OF 1-349 FROM N.A.
RC TISSUE=Brain;
RA Ikeda M., Takehara N., Ebisawa T., Yamauchi T., Nomura M.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBSJ databases.
CC !- FUNCTION: CIRCADIAN REGULATOR THAT ACTS AS A TRANSCRIPTION FACTOR.
CC CLOCK-BMAL1 HETERODIMERS BIND TO AN E-BOX ELEMENT (3'-CACGTG-5'),
CC THEREBY ACTIVATING TRANSCRIPTION OF PER1, AND POSSIBLY OF OTHER
CC CIRCADIAN CLOCK PROTEINS. MUTANT CLOCK AND BMAL1 FORM HETERODIMER
CC THAT BIND DNA, BUT FAIL TO ACTIVATE TRANSCRIPTION (BY SIMILARITY).
CC !- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN. HETERODIMERS WITH BMAL1, AND LESS EFFICIENTLY WITH
CC ARNT AND ARNT2. HETERODIMERS WITH ARNT OR ARNT2 BIND POORLY TO THE
CC E-BOX MOTIF (BY SIMILARITY).
CC !- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC !- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC !- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; AF011568; AAB83969.1; -.
DR EMBL; AF097458; AAF13733.1; -.
DR EMBL; AF097442; AAF13733.1; JOINED.
DR EMBL; AF097443; AAF13733.1; JOINED.
DR EMBL; AF097444; AAF13733.1; JOINED.
DR EMBL; AF097445; AAF13733.1; JOINED.
DR EMBL; AF097446; AAF13733.1; JOINED.
DR EMBL; AF097447; AAF13733.1; JOINED.
DR EMBL; AF097448; AAF13733.1; JOINED.
DR EMBL; AF097449; AAF13733.1; JOINED.
DR EMBL; AF097450; AAF13733.1; JOINED.
DR EMBL; AF097451; AAF13733.1; JOINED.
DR EMBL; AF097452; AAF13733.1; JOINED.
DR EMBL; AF097453; AAF13733.1; JOINED.
DR EMBL; AF097454; AAF13733.1; JOINED.
DR EMBL; AF097455; AAF13733.1; JOINED.
DR EMBL; AF097456; AAF13733.1; JOINED.
DR EMBL; AF097457; AAF13733.1; JOINED.
DR EMBL; AB002332; BAA20792.1; -.
DR EMBL; AB005535; BAA21774.1; -.
DR MIM; 601851; -.
DR InterPro; IPR003015; HLH_MYC.
DR InterPro; IPR001092; HLH_dlm.
DR InterPro; IPR001067; Nucleinslocatr.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR Pfam; PF00785; PAC; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR00785; NCTRNLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
KW Transcription regulation; Nuclear protein; Repeat; Biological rhythms;
KW DNA-binding.
FT DNA_BIND 35 47
FT DOMAIN 48 85
FT REPEAT 109 175
FT REPEAT 264 330
FT DOMAIN 514 564
FT DOMAIN 744 760
FT DOMAIN 819 828
FT CONFLICT 440 440
S -> P (IN REF. 2).
```

SQ SEQUENCE 846 AA; 95303 MW; C292B451A33E4CBF CRC64;

Query Match 54.9%; Score 39; DB 1; Length 846;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
| |||||  
Db 344 KGQOWIWL 351

RESULT 26  
CLOC\_MOUSE STANDARD: PRT; 855 AA.  
AC O08785;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE CIRCADIAN LOCOMOTOR OUTPUT CYCLES KAPUT PROTEIN (MCLOCK).  
GN CLOC.

OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129;  
RX MEDLINE=97304393; PubMed=9160756;  
RA Antoch M.P., Song E.J., Chang A.M., Vitaterna M.H., Zhao Y.,  
Wilsbacher L.D., Sangoram A.M., King D.P., Pinto L.H., Takahashi J.S.;  
RT "Functional identification of the mouse circadian clock gene by  
transgenic BAC rescue";  
RL Cell 89:655-667(1997).  
RN [2]  
RN SEQUENCE FROM N.A., AND IDENTIFICATION OF CLOCK MUTANT.  
RC STRAIN=BALB/C X C57BL/6; Tissue-Suprachiasmatic nucleus;  
RX MEDLINE=97304392; PubMed=9160755;  
RA King D.P., Zhao Y., Sangoram A.M., Wilsbacher L.D., Tanaka M.,  
Antoch M.P., Steeves J.D.L., Vitaterna M.H., Kornhauser J.M.,  
Lowrey P.L., Turek F.W., Takahashi J.S.;  
RT "Positional cloning of the mouse circadian clock gene";  
RL Cell 89:641-653(1997).  
RN [3]  
RN SEQUENCE FROM N.A.  
RC STRAIN=129/SV;  
RA Wilsbacher L.D., Sangoram A.M., Antoch M.P., Takahashi J.S.;  
RT "The mouse clock locus: Sequence and analysis of 204 kb from mouse  
chromosome 5";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RN INTERACTION WITH BMAL1.  
RX MEDLINE=98279137; PubMed=9616112;  
RA Gekakis N., Staknis D., Nguyen H.B., Davis F.C., Wilsbacher L.D.,  
King D.P., Takahashi J.S., Weitz C.J.;  
RT "Role of the clock protein in the mammalian circadian mechanism";  
RL Science 280:1564-1569(1998).

CC -1- FUNCTION: CIRCADIAN REGULATOR THAT ACTS AS A TRANSCRIPTION FACTOR.  
CC CLOCK-BMAL1 HETERODIMERS BIND TO AN E-BOX ELEMENT (3'-CACGTG-5'),  
CC THEREBY ACTIVATING TRANSCRIPTION OF PER1, AND POSSIBLY OF OTHER  
CC CIRCADIAN CLOCK PROTEINS. MUTANT CLOCK AND BMAL1 FORM HETERODIMER  
CC THAT BIND DNA, BUT FAIL TO ACTIVATE TRANSCRIPTION. IN HOMOZYGOUS  
CC CLOCK MUTANTS, THE CIRCADIAN PERIOD IS INCREASED FROM 3 TO 4 HOURS  
CC AND USUALLY THE CIRCADIAN RHYTHMICITY IS LOST AT CONSTANT  
CC DARKNESS. EXPRESSION OF CLOCK IS ALSO REDUCED.  
CC -1- SUBUNIT: HETERODIMER WITH BMAL1, AND LESS EFFICIENTLY WITH ARNT  
CC AND ARNT2. HETERODIMERS WITH ARNT OR ARNT2 BIND POORLY TO THE E-  
CC BOX MOTIF.  
CC -1- TISSUE SPECIFICITY: EXPRESSED EQUALLY IN BRAIN, EYE, TESTES,  
CC OVARIES, LIVER, HEART, LUNG, KIDNEY, IN THE BRAIN, EXPRESSION IS  
CC ABUNDANT IN THE SUPRACHIASMATIC NUCLEI (SCN), IN THE PYRIFORM  
CC CORTEX, AND IN THE HIPPOCAMPUS. LOW EXPRESSION THROUGHOUT THE REST  
CC OF THE BRAIN. EXPRESSION DOES NOT APPEAR TO UNDERGO CIRCADIAN

CC OSCILLATIONS.  
CC -1- DOMAIN: CONTAINS A GLN-RICH C-TERMINAL DOMAIN WHICH COULD  
CC CORRESPOND TO THE TRANSACTIVATION DOMAIN. IN MUTANT CLOCK,  
CC DELETION OF THIS REGION LEADS TO AN INCREASED CIRCADIAN PERIOD  
CC FROM 3 TO 4 HOURS AS WELL AS TO THE LOSS OF CIRCADIAN RHYTHMICITY  
CC AND ALTERED LIGHT RESPONSE.  
CC -1- DISEASE: DEFECTS IN CLOCK AFFECT TWO PROPERTIES OF THE CIRCADIAN  
CC SYSTEM: THE LENGTH OF THE FREE-RUNNING PERIOD AND THE PERSISTENCE  
CC OF CIRCADIAN RHYTHMICITY IN CONSTANT DARKNESS.  
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (3HLH) FAMILY OF  
CC TRANSCRIPTION FACTORS.  
CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF000998; AAC53200.1; -;  
CC EMBL; AF146793; AAD30565.1; -;  
CC MG1; 99698; CLOC.  
CC InterPro: IPR003015; HLH\_MYC.  
CC InterPro: IPR001092; HLH\_dim.  
CC InterPro: IPR001067; Nucleoslocatr.  
CC InterPro: IPR001610; PAC.  
CC InterPro: IPR000014; PAS.  
CC InterPro: IPR003617; TFS2\_N.  
CC Pfam: PF00785; PAC; 1.  
CC Pfam: PF00989; PAC; 2.  
CC PRINTS; PR00785; NCTRNLOCATR.  
CC SMART; SM00353; HLH; 1.  
CC SMART; SM00086; PAC; 1.  
CC SMART; SM00509; TFS2N; 1.  
CC SMART; SM00091; PAS; 2.  
CC PROSITE; PS00038; HELIX\_LOOP\_HELIX; 1.  
KW Transcription regulation; Nuclear protein; Repeat; Biological rhythms;  
KW DNA-binding; Alternative splicing.  
FT DNA\_BIND 35 47 BASIC DOMAIN.  
FT DOMAIN 48 85 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).  
FT REPEAT 109 175 PAS-1.  
FT REPEAT 264 330 PAS-2.  
FT DOMAIN 484 855 GLN-RICH.  
FT DOMAIN 740 745 POLY-GLN.  
FT DOMAIN 751 759 POLY-GLN.  
FT DOMAIN 762 769 POLY-GLN.  
FT DOMAIN 828 837 POLY-GLN.  
FT DOMAIN 514 564 IMPLICATED IN THE CIRCADIAN RHYTHMICITY.  
FT VARSPIC 484 513 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 855 AA; 96393 MW; 9864D947049742F4 CRC64;

Query Match 54.9%; Score 39; DB 1; Length 855;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
| |||||  
Db 344 KGQOWIWL 351

RESULT 27  
CLOC\_MOUSE STANDARD: PRT; 1023 AA.  
ID CLOC\_MOUSE  
AC O61735; O76342; O77137;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CIRCADIAN LOCOMOTOR OUTPUT CYCLES KAPUT PROTEIN (DCLOCK) (DPAS1).  
GN CLK OR JRK OR CLOCK OR PAS1.  
OS Drosophila melanogaster (Fruit fly).

CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Head;  
RX MEDLINE=98279147; PubMed=9616122;  
RA Darlington T.K., Wager-Smith K., Ceriani M.F., Staknis D., Gekakis N.,  
RA Steeves T.D.L., Weitz C.J., Takahashi J.S., Kay S.A.;  
RT "Closing the circadian loop: CLOCK-induced transcription of its own  
RT inhibitors per and tim.";  
RL Science 280:1599-1603(1998).  
[2]  
RP SEQUENCE FROM N.A., AND MUTAGENESIS.  
RC TISSUE=Head;  
RX MEDLINE=98292177; PubMed=9630223;  
RA Allada R., White N.E., So W.V., Hall J.C., Rosbash M.;  
RT "A mutant Drosophila homolog of mammalian Clock disrupts circadian  
RT rhythms and transcription of period and timeless.";  
RL Cell 93:791-804(1998).  
[3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CANTON-S;  
RX MEDLINE=98414630; PubMed=9742131;  
RA Bae K., Lee C., Sidote D., Chuang K.-Y., Edery I.;  
RT "Circadian regulation of a Drosophila homolog of the mammalian clock  
RT gene: PER and TIM function as positive regulators.";  
RL Mol. Cell. Biol. 18:6142-6151(1998).  
CC -1- FUNCTION: CIRCADIAN REGULATOR THAT ACTS AS A TRANSCRIPTION FACTOR  
CC AND GENERATES A RHYTHMIC OUTPUT WITH A PERIOD OF ABOUT 24 HOURS.  
CC OSCILLATES IN ANTIPHASE TO THE CYCLING OBSERVED FOR PERIOD (PER)  
CC AND TIMELESS (TIM). ACCORDING TO REF.3, REACHES PEAK ABUNDANCE  
CC WITHIN SEVERAL HOURS OF THE DARK-LIGHT TRANSITION AT ZT0  
CC (ZEITGEBER 0), WHEREAS REF.1 DESCRIBES BIMODAL OSCILLATING  
CC EXPRESSION WITH MAXIMUM AT ZT5 AND ZT23. CLOCK-CYCLE HETERODIMERS  
CC ACTIVATE CYCLING TRANSCRIPTION OF PER AND TIM BY BINDING TO THE E-  
CC BOX (3'-CACGTG-5') PRESENT IN THEIR PROMOTERS. ONCE INDUCED,  
CC PERIOD AND TIMELESS BLOCK CLOCK'S ABILITY TO TRANSDUCE THEIR  
CC PROMOTERS.  
CC -1- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER  
CC BHLH PROTEIN. FORMS A HETERODIMER WITH CYCLE.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; THE FULL-LENGTH VARIANT A  
CC (SHOWN HERE) AND VARIANT B; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC VARIANT B ENCODES TWO CONCEPTUAL PROTEINS. THE FIRST CONSISTS ONLY  
CC OF THE BHLH DOMAIN, THE OTHER CONSISTS OF THE PAS-1 AND ALL C-  
CC TERMINAL DOMAINS. VARIANT B IS EXPRESSED WEAKLY AT ALL THE TIME OF  
CC THE DAY, AND IT CYCLES IN PHASE WITH THE FULL-LENGTH FORM.  
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED. FOUND IN HEAD, BODY, AND  
CC APPENDAGE FRACTIONS.  
CC -1- DOMAIN: CONTAINS THREE POLYGLUTAMINE REPEATS WHICH COULD  
CC CORRESPOND TO THE TRANSDUCATION DOMAIN. THE LENGTH OF THE  
CC REPEATS IS POLYMORPHIC. IN THE ARRYTHMIC MUTANT JRK, DELETION OF  
CC THIS REGION LEADS TO THE LOSS OF CIRCADIAN RHYTHMICITY AND ALTERED  
CC LIGHT RESPONSE.  
CC -1- POLYMORPHISM: THE VARIABILITY IN LENGTH OF THE POLYGLUTAMINE  
CC STRETCH IS DUE TO POLYMORPHISM OF THIS REGION.  
CC -1- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF  
CC TRANSCRIPTION FACTORS.  
CC -1- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF067207; AAD10630.1; -  
DR EMBL; AF065133; AAC39101.1; -  
DR EMBL; AF069997; AAC62234.1; -

DR FlyBase; FBgn0023076; clk.  
DR InterPro; IPR003015; HLH\_Myc.  
DR InterPro; IPR001092; HLH\_dim.  
DR InterPro; IPR001067; Nucleotranslocatr.  
DR InterPro; IPR001610; PAC.  
DR InterPro; IPR000014; PAC.  
DR Pfam; PF00010; HLH; 1.  
DR Pfam; PF00785; PAC; 1.  
DR Pfam; PF00989; PAS; 2.  
DR PRINTS; PR00785; NCTRNLOCATR.  
DR SMART; SM00353; HLH; 1.  
DR SMART; SM00086; PAC; 1.  
DR SMART; SM00091; PAS; 2.  
DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; 1.  
KW Transcription regulation; Nuclear protein; Repeat; Biological rhythms;  
KW DNA-binding; Alternative splicing.  
FT DNA\_BIND 12 24  
FT DOMAIN 25 62  
FT REPEAT 86 152  
FT REPEAT 251 317  
FT DOMAIN 548 559  
FT DOMAIN 766 769  
FT DOMAIN 794 836  
FT DOMAIN 874 877  
FT DOMAIN 887 895  
FT DOMAIN 953 963  
FT DOMAIN 776 1023  
FT VARIANT 816 823  
FT CONFLICT 12 12  
FT CONFLICT 32 32  
FT CONFLICT 128 128  
FT CONFLICT 555 555  
FT CONFLICT 605 605  
FT CONFLICT 912 912  
SQ SEQUENCE 1023 AA; 115751 MW; 514374CBC050DRAFB CRC64;  
  
Query Match 54.9%; Score 39; DB 1; Length 1023;  
Best Local Similarity 75.0%; Pred. No. 1.8e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 3 KPQQFWL 10  
DB 331 KGQQWIL 338  
  
RESULT 28  
MSBB\_ECOLI ID MSBB\_ECOLI STANDARD; PRT; 323 AA.  
AC P24205;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DE 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LIPID A BIOSYNTHESIS (KDO)2-(LAUROYL)-LIPID IVA ACYLTRANSFERASE  
DE (EC 2.3.1.-)  
GN MSBB OR B1855.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / W3110;  
RX MEDLINE=92121107; PubMed=1732206;  
RA Karow M., Georgopoulos C.;  
RT "Isolation and characterization of the Escherichia coli msbB gene, a  
RT multicopy suppressor of null mutations in the high-temperature  
RT requirement gene htrB".  
RL J. Bacteriol. 174:702-710(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=93015688; PubMed=1356966;





```
FT MOD_RES 77 77 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT MOD_RES 78 78 PHOSPHORYLATION (BY CK2) (POTENTIAL).
FT VARSPLIC 1 23 MAERAIIDPCQEDELHAEDEG -> MNGFALLRRNAS
FT CONFLICT 155 156 KRGLKLLR (IN ISOFORM S37).
FT CONFLICT 160 160 AE -> PK (IN REF. 1).
FT CONFLICT 160 160 I -> T (IN REF. 1).
FT CONFLICT 168 168 S -> N (IN REF. 1).
FT CONFLICT 253 253 S -> G (IN REF. 1).
SQ SEQUENCE 330 AA; 36714 MW; CCC27150F02859FB CRC64;

Query Match 53.5%; Score 38; DB 1; Length 330;
Best Local Similarity 55.6%; Pred. No. 90;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQQFW 9
Db 92 KPEPRQQFW 100

RESULT 30
FUT4_HUMAN
AC P22083;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE ALPHA-(1,3)-FUCOSYLTRANSFERASE (EC 2.4.1.-) (GALACTOSIDE 3-L-
DE FUCOSYLTRANSFERASE) (FUCOSYLTRANSFERASE 4) (FUCT-IV) (ELAM-1 LIGAND
DE FUCOSYLTRANSFERASE).
GN FUT4 OR ELFT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Peripheral blood leukocytes;
RX MEDLINE=91373370; PubMed=1716630;
RA Lowe J.B., Kukowska-Latallo J.F., Nair R.P., Larsen R.D., Marks R.M.,
RA Macher B.A., Kelly R.J., Ernst L.K.;
RT "Molecular cloning of a human fucosyltransferase gene that determines
RT expression of the Lewis x and VIM-2 epitopes but not ELAM-1-dependent
RT cell adhesion.";
RL J. Biol. Chem. 266:17467-17477(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91084863; PubMed=1702034;
RA Goelz S.E., Hession C., Goff D., Griffiths B., Tizard R., Newman B.,
RA Chi-Rosso G., Lobb R.;
RT "ELFT: a gene that directs the expression of an ELAM-1 ligand.";
RL Cell 63:1349-1356(1990).
RN [3]
RP SEQUENCE OF 1-400 FROM N.A.
RX MEDLINE=92042084; PubMed=1718983;
RA Kumar R., Potvin B., Muller W.A., Stanley P.;
RT "Cloning of a human alpha(1,3)-fucosyltransferase gene that encodes
RT ELFT but does not confer ELAM-1 recognition on Chinese hamster ovary
RT cell transfectants.";
RL J. Biol. Chem. 266:21777-21783(1991).
CC -1- FUNCTION: MAY CATALYSE ALPHA-1,3 GLYCOSIDIC LINKAGES INVOLVED IN
CC THE EXPRESSION OF LEWIS X/SSEA-1 AND VIN-2 ANTIGENS.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC
DR EMBL: M65030; AAA92977.1; -
DR EMBL: M58596; AAA63172.1; -
DR EMBL: M58597; AAA63173.1; ALT_INIT.
DR EMBL: S65161; AAB20349.1; -
DR PIR: A36340; A36340.
DR MIM: 104230; -
DR InterPro: IPR001503; Glyco.transf_10.
DR Pfam: PF00852; Glyco.transf_10; 1.
KW transferase; Glycosyltransferase; Transmembrane; Glycoprotein;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 23 47 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT DOMAIN 48 405 (POTENTIAL).
FT CARBOHYD 91 91 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 190 190 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 87 87 P -> R (IN REF. 2 AND 3).
FT CONFLICT 241 241 E -> D (IN REF. 3).
SQ SEQUENCE 405 AA; 45569 MW; DE72E1FDC390268D CRC64;

Query Match 53.5%; Score 38; DB 1; Length 405;
Best Local Similarity 50.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQQFW 10
Db 161 RPPGQRWVM 170

RESULT 31
FUT4_MOUSE
ID FUT4_MOUSE STANDARD; PRT; 433 AA.
AC Q11127;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ALPHA-(1,3)-FUCOSYLTRANSFERASE (EC 2.4.1.-) (GALACTOSIDE 3-L-
DE FUCOSYLTRANSFERASE) (FUCOSYLTRANSFERASE 4) (FUCT-IV).
GN FUT4 OR ELFT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96027607; PubMed=7559635;
RA Gersten K.M., Natsuka S., Trinchera M., Petryniak B., Kelly R.J.,
RA Hiraawa N., Jenkins N.A., Gilbert D.J., Copeland N.G., Lowe J.B.;
RT "Molecular cloning, expression, chromosomal assignment, and tissue-
RT specific expression of a murine alpha-(1,3)-fucosyltransferase locus
RT corresponding to the human ELAM-1 ligand fucosyl transferase.";
RL J. Biol. Chem. 270:25047-25056(1995).
RN [2]
RP SEQUENCE FROM N.A. (SHORT FORM).
RC STRAIN=129/SV; TISSUE=Liver;
RX MEDLINE=97037075; PubMed=8882722;
RA Ozawa M., Muramatsu T.;
RT "Molecular cloning and expression of a mouse alpha-1,3
RT fucosyltransferase gene that shows homology with the human alpha-1,3
RT fucosyltransferase IV gene";
RL J. Biochem. 119:302-308(1996).
CC -1- FUNCTION: MAY CATALYSE ALPHA-1,3 GLYCOSIDIC LINKAGES INVOLVED IN
CC THE EXPRESSION OF LEWIS X/SSEA-1 AND VIN-2 ANTIGENS.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS. A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN STOMACH AND COLON.
CC IT ALSO EXPRESSED IN THE LUNG, TESTIS, UTERUS, SMALL INTESTINE
CC AND TO A LESSER EXTENT IN SPLEEN, AND OVARY. PRESENT IN TRACE
```

CC AMOUNTS IN BRAIN, THYMUS, HEART, SMOOTH MUSCLE, KIDNEY AND BONE  
CC MARROW. NOT FOUND IN LIVER, SALIVARY GLAND AND PANCREAS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U33457; AAC5269.1; -  
CC DR EMBL: D63380; BAA09697.1; -  
CC DR EMBL: D63379; BAA09696.1; -  
CC DR MGI: 95594; Fut4.  
CC DR InterPro: IPR001503; Glyco\_transf\_10.  
CC DR Pfam: PF00852; Glyco\_transf\_10; 1.  
CC KW Transferase; Glycosyltransferase; Transmembrane; Glycoprotein;  
CC Signal-anchor; Golgi stack; Alternative splicing.  
CC DOMAIN 1 52 CYTOPLASMIC (POTENTIAL).  
CC FT TRANSMEM 53 74 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
CC (POTENTIAL).  
CC FT DOMAIN 75 433 LUMENAL, CATALYTIC (POTENTIAL).  
CC FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT VARSPLIC 218 218 MISSING (IN SHORT ISOFORM).  
CC FT CONFLICT 252 252 Q -> P (IN REF. 2).  
CC FT CONFLICT 257 257 R -> Q (IN REF. 2).  
CC FT CONFLICT 260 260 V -> E (IN REF. 2).  
CC FT CONFLICT 273 273 R -> Q (IN REF. 2).  
CC SQ SEQUENCE 433 AA; 49481 MW; 2401822F02B5D021 CRC64;

Query Match 53.5%; Score 38; DB 1; Length 433;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
||| | | |  
DB 189 RPPGQRVWM 198

RESULT 32  
FUT4\_RAT  
ID FUT4\_RAT STANDARD; PRT; 433 AA.  
AC Q62994;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE ALPHA-(1.3)-FUCOSYLTRANSFERASE (EC 2.4.1.-) (GALACTOSIDE 3-L-  
DE FUCOSYLTRANSFERASE) (FUCOSYLTRANSFERASE 4) (FUCT-IV).  
GN FUT4 OR RFUC-T.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OC NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Kidney;  
RX MEDLINE=97265205; PubMed=911142;  
RA Szejdel-Sulkowska E.M., Smith F.I., Wiedersheim G., McCluer R.H.;  
RT "Cloning of a rat alpha1,3-fucosyltransferase gene: a member of the  
RT fucosyltransferase IV family.";  
RL Glycoconj. J. 14:249-258(1997).  
CC -1- FUNCTION: MAY CATALYSE ALPHA-1,3 GLYCOSIDIC LINKAGES INVOLVED IN  
CC THE EXPRESSION OF LEWIS X/SSA-1 AND VIM-2 ANTIGENS.  
CC -1- PATHWAY: GLYCOSYLATION.  
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
CC FORM IN TRANS CISTERNAE OF GOLGI.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: IN ADULT, HIGHEST EXPRESSION IN SPLEEN,  
CC TESTIS, BRAIN, LUNG, KIDNEY AND SKELETAL MUSCLE AND TO A LESSER

CC EXTENT IN LIVER AND HEART.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U58860; AAB97609.1; -  
CC DR InterPro: IPR001503; Glyco\_transf\_10.  
CC DR Pfam: PF00852; Glyco\_transf\_10; 1.  
CC KW Transferase; Glycosyltransferase; Transmembrane; Glycoprotein;  
CC Signal-anchor; Golgi stack; Alternative splicing.  
CC DOMAIN 1 54 CYTOPLASMIC (POTENTIAL).  
CC FT TRANSMEM 55 74 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
CC (POTENTIAL).  
CC FT DOMAIN 75 433 LUMENAL, CATALYTIC (POTENTIAL).  
CC FT CARBOHYD 117 117 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT VARSPLIC 218 218 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT VARSPLIC 33 MISSING (IN SHORT ISOFORM).  
CC SQ SEQUENCE 433 AA; 48779 MW; 75B0E569B72FD2F8 CRC64;  
Query Match 53.5%; Score 38; DB 1; Length 433;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
||| | | |  
DB 189 RPPGQRVWM 198

RESULT 33  
YADC\_SCHPO  
ID YADC\_SCHPO STANDARD; PRT; 533 AA.  
AC Q09837;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE HYPOTHETICAL 62.2 KDA PROTEIN C4G8.12C IN CHROMOSOME I.  
GN SPAC4G8.12C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA Badcock K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: TO YEAST SMP3.  
CC -1- SIMILARITY: SOME, TO YEAST YGL142C.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z56276; CAA91213.1; -  
CC DR Hypothetical protein; Transmembrane.  
CC KW TRANSMEM 8 28 POTENTIAL.  
CC FT TRANSMEM 61 81 POTENTIAL.  
CC FT TRANSMEM 91 111 POTENTIAL.  
CC FT TRANSMEM 144 164 POTENTIAL.  
CC FT TRANSMEM 175 195 POTENTIAL.  
CC FT TRANSMEM 216 236 POTENTIAL.

```
FT TRANSMEM 274 294 POTENTIAL.
FT TRANSMEM 297 317 POTENTIAL.
FT TRANSMEM 338 358 POTENTIAL.
FT TRANSMEM 496 516 POTENTIAL.
SQ SEQUENCE 533 AA; 62200 MW; F14519C95884687 CRC64;

Query Match 53.5%; Score 38; DB 1; Length 533;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 RPKQWFWLM 11
   ||||
Db 296 KPAWLL 304

RESULT 34
ARNT_HUMAN
ID ARNT_HUMAN STANDARD; PRT; 789 AA.
AC P27540;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ARYL HYDROCARBON RECEPTOR NUCLEAR TRANSLOCATOR (ARNT PROTEIN) (DIOXIN
DE RECEPTOR, NUCLEAR TRANSLOCATOR) (HYPOXIA-INDUCIBLE FACTOR 1 BETA)
DE (HIF-1 BETA).
GN ARNT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. PubMed=1852076;
RX MEDLINE=91240280; PubMed=1852076;
RA Hoffman E.C., Reyes H., Chu F.-F., Sander F., Conley L.H.,
RA Brooks B.A., Hankinson O.;
RT "Cloning of a factor required for activity of the Ah (dioxin)
RT receptor.";
RL Science 252:954-958(1991).
RN [2]
RP SEQUENCE OF 186-203 AND 662-694.
RX MEDLINE=95296340; PubMed=7539918;
RA Wang G.L., Jiang B.-H., Rue E.A., Semenza G.L.;
RT "Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS
RT heterodimer regulated by cellular O2 tension.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:5510-5514(1995).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=92271249; PubMed=1317062;
RA Reyes H., Reisz-Porszasz S., Hankinson O.;
RT "Identification of the Ah receptor nuclear translocator protein
RT (Arnt) as a component of the DNA binding form of the Ah receptor.";
RL Science 256:1193-1195(1992).
CC -!- FUNCTION: REQUIRED FOR ACTIVITY OF THE AH (DIOXIN) RECEPTOR. THIS
CC PROTEIN IS REQUIRED FOR THE LIGAND-BINDING SUBUNIT TO TRANSLOCATE
CC FROM THE CYTOSOL TO THE NUCLEUS AFTER LIGAND BINDING. THE COMPLEX
CC THEN INITIATES TRANSCRIPTION OF A GENES INVOLVED IN THE ACTIVATION
CC OF PAH PROCARCINOGENS.
CC -!- SUBUNIT: EFFICIENT DNA BINDING REQUIRES DIMERIZATION WITH ANOTHER
CC BHLH PROTEIN. FORMS AN HETERODIMER WITH AHR, WITH HIF1A AS WELL AS
CC WITH OTHER BHLH PROTEINS. INTERACTS WITH TRANSFORMING ACIDIC
CC COILED-COIL CONTAINING PROTEIN 3 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND
CC A SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE BASIC HELIX-LOOP-HELIX (BHLH) FAMILY OF
CC TRANSCRIPTION FACTORS.
CC -!- SIMILARITY: CONTAINS 1 PAS (PER-ARNT-SIM) DIMERIZATION DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M69238; AAA51777.1; -.
DR HSSP; P22415; IANA.
DR TRANSFAC; T01346; -.
DR MIM; 126110; -.
DR InterPro; IPR003015; HLH_Myc.
DR InterPro; IPR001092; HLH_dim.
DR InterPro; IPR001067; Nuctrnslocatr.
DR InterPro; IPR001610; PAC.
DR InterPro; IPR000014; PAS.
DR Pfam; PF00010; HLH; 1.
DR Pfam; PF00989; PAS; 2.
DR PRINTS; PR00785; NCTRNSLOCATR.
DR SMART; SM00353; HLH; 1.
DR SMART; SM00086; PAC; 1.
DR SMART; SM00091; PAS; 2.
DR PROSITE; PS00038; HELIX_LOOP_HELIX; 1.
KW Nuclear protein; DNA-binding; Transcription regulation; Activator;
KW Alternative splicing; Repeat.
FT DNA_BIND 90 102 BASIC DOMAIN.
FT DOMAIN 103 143 HELIX-LOOP-HELIX MOTIF (BY SIMILARITY).
FT REPEAT 163 230 PAS-1.
FT REPEAT 351 417 PAS-2.
FT DOMAIN 424 467 PAC MOTIF.
FT DOMAIN 710 769 GLN-RICH.
FT DOMAIN 99 102 POLY-ARG.
FT DOMAIN 503 507 POLY-GLN.
FT DOMAIN 738 741 POLY-SER.
FT VARSPLOC 77 91 MISSING (IN SHORT ISOFORM).
SQ SEQUENCE 789 AA; 86636 MW; 2E278F8E62BFBF6D CRC64;

Query Match 53.5%; Score 38; DB 1; Length 789;
Best Local Similarity 50.0%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKQWFWLM 10
   |||||
Db 430 RSKNQEWLWM 439

RESULT 35
VEF_GVPV
ID VEF_GVPV STANDARD; PRT; 901 AA.
AC P41723;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE VIRAL ENHANCING FACTOR (VEF) (ENHANCIN) (104 KDA GLYCOPROTEIN)
DE (SYNERGISTIC FACTOR).
GN VEF.
OS Pseudolatia unipuncta granulosis virus (puGV) (Pseudolatia unipuncta
OS granulovirus).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=36355;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HAWAII;
RX MEDLINE=96068802; PubMed=7595376;
RA Roelvink P.W., Corsaro B.G., Granados R.R.;
RT "Characterization of the Helicoverpa armigera and Pseudolatia
RT unipuncta granulovirus enhancin genes.";
RL J. Gen. Virol. 76:2693-2705(1995).
CC -!- FUNCTION: INVOLVED IN DISRUPTION OF THE PERITROPHIC MEMBRANE AND
CC FUSION OF NUCLEOCAPSIDS WITH MIDGUT CELLS (BY SIMILARITY).
CC -!- SIMILARITY: TO TNGV AND HAGV VEF.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; D14871; BAA03587.1; --  
KW Glycoprotein; Late protein.  
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 349 349 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 642 642 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 683 683 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 698 698 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 901 AA; 104252 MW; 5D0E542A858FE5FB CRC64;

Query Match 53.5%; Score 38; DB 1; Length 901;  
Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQOWFWL 10  
DB 353 PYQIWSWL 361  
I I I I I

RESULT 36  
VEF\_GVTN STANDARD; PRT; 901 AA.  
ID VEF\_GVTN  
AC P29998;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE VIRAL ENHANCING FACTOR (VEF) (ENHANCIN) (104 KDA GLYCOPROTEIN)  
GN (SYNERGISTIC FACTOR).  
DE VEF.

OS Trichoplusia ni granulosis virus (TnGV) (Trichoplusia ni  
OS granulovirus).  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.  
OX NCBI\_TaxID=10462;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=92044434; PubMed=1940861;  
RA Hashimoto Y., Corsaro B.G., Granados R.R.;  
RT "Location and nucleotide sequence of the gene encoding the viral  
RT enhancing factor of the Trichoplusia ni granulosis virus.";  
RL J. Gen. Virol. 72:2645-2651(1991).  
CC -|- FUNCTION: INVOLVED IN DISRUPTION OF THE PERITROPHIC MEMBRANE AND  
CC FUSION OF NUCLEOCAPSIDS WITH MIDGUT CELLS.  
CC -|- SIMILARITY: TO PUGV AND HAGV VEF.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; D12617; BAA02141.1; --  
DR PIR; J01328; WNVNVT.  
KW Glycoprotein; Late protein.  
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 349 349 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 642 642 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 683 683 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 698 698 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 901 AA; 104322 MW; 74DB822E0A11CD6A CRC64;

Query Match 53.5%; Score 38; DB 1; Length 901;  
Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQOWFWL 10  
DB 353 PYQIWSWL 361  
I I I I I

RESULT 37  
Y812\_ARCFU STANDARD; PRT; 298 AA.  
ID Y812\_ARCFU  
AC Q29446;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL PROTEIN AF0812.  
GN AF0812.

OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
OC Archaeoglobus.  
OX NCBI\_TaxID=2234;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE=98049343; PubMed=9389475;  
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodak A., Zhou L.,  
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,  
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;

RT "The complete genome sequence of the hyperthermophilic, sulphate-  
RT reducing archaeon Archaeoglobus fulgidus.";  
RL Nature 390:364-370(1997).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AE001048; AAB90432.1; --  
DR TIGR; AF0812; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 298 AA; 34385 MW; ED59E86A07AC5A30 CRC64;

Query Match 52.8%; Score 37.5; DB 1; Length 298;  
Best Local Similarity 66.7%; Pred. No. 97;  
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 2 PKPQOWFW 9  
DB 62 PKPEYFW 70  
I I I I I

RESULT 38  
DNBI\_HSV1 STANDARD; PRT; 375 AA.  
ID DNBI\_HSV1  
AC Q03444;

DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DE 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MAJOR DNA-BINDING PROTEIN (FRAGMENT).  
GN 31.  
OS Equine herpesvirus type 1 (isolate HVS25A) (EHV-1).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=10327;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94106109; PubMed=8279122;  
RA Bell C.W., Whalley J.M.;  
RA "Herpesvirus ICP18.5 and DNA-binding protein genes are conserved in  
RT equine herpesvirus-1";  
RT equine herpesvirus-1";  
RL Virus Genes 7:219-228(1993).  
CC -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA  
CC REPLICATION.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D13930; BAA03033.1; -;  
DR PIR; JQ0846; JQ0846.  
KW DNA-binding; DNA replication; Nuclear protein.  
FT NON\_TER 1  
SQ SEQUENCE 375 AA; 40309 MW; ECF327925EBF999B CRC64;  
  
Query Match 52.8%; Score 37.5; DB 1; Length 375;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
  
QY 2 PRKPOQFWLM 11  
Db 10 PNP-QWFWTL 18  
: | | | | |  
P N P Q W F W T L 1 8  
  
RESULT 39  
ID DNBI\_HSV11 STANDARD; PRT; 1196 AA.  
AC P04296;  
DT 20-MAR-1987 (Rel. 04, Created)  
DT 20-MAR-1987 (Rel. 04, Last sequence update)  
DE MAJOR DNA-BINDING PROTEIN (INFECTED CELL PROTEIN 8) (ICP 8 PROTEIN).  
GN DBP OR UL29 OR ICP8.  
OS Herpes simplex virus (type 1 / strain 17).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88274327; PubMed=2839594;  
RA McGeech D.J., Dairymple M.A., Davison A.J., Dolan A., Frame M.C.,  
RA McNab D., Perry L.J., Scott J.E., Taylor P.;  
RT "The complete DNA sequence of the long unique region in the genome of  
RT herpes simplex virus type 1";  
RL J. Gen. Virol. 69:1531-1574(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86067223; PubMed=2999714;  
RA Quinn J.P., McGeoch D.J.;  
RT "DNA sequence of the region in the genome of herpes simplex virus  
RT type 1 containing the genes for DNA polymerase and the major DNA

binding protein.";  
RL Nucleic Acids Res. 13:8143-8163(1985).  
RN [3]  
RP SEQUENCE OF 1062-1196 FROM N.A.  
RX MEDLINE=88306232; PubMed=2457278;  
RA Hammerschmidt W., Conraths F., Mankertz J., Buhk H.-J., Pauli G.,  
RA Ludwig H.;  
RT "Common epitopes of glycoprotein B map within the major DNA-binding  
RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex  
RT virus type 1 (HSV-1).";  
RL Virology 165:406-418(1988).  
CC -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA  
CC REPLICATION.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR. IN THE ABSENCE OF DNA REPLICATION,  
CC FOUND IN THE NUCLEAR FRAMEWORK-ASSOCIATED STRUCTURES  
CC (PEREPLICATIVE SITES); AS VIRAL DNA REPLICATION PROCEEDS, IT  
CC MIGRATES TO GLOBULAR INTRANUCLEAR STRUCTURES (REPLICATION  
CC COMPARTMENTS).  
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN  
CC FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D10879; BAA01675.1; -;  
DR EMBL; X14112; CAA32322.1; -;  
DR EMBL; X03181; CAA26940.1; -;  
DR EMBL; M21631; AAA45787.1; -;  
DR PIR; A03790; DNBEV1.  
DR PIR; B30085; B30085.  
DR InterPro; IPR000635; Viral\_DNA\_bind.  
DR Pfam; PF00747; viral\_DNA\_bp.1.  
KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.  
FT ZN\_FING 499 512 C2HC-TYPE  
SQ SEQUENCE 1196 AA; 128349 MW; 4537991625B99E9 CRC64;  
  
Query Match 52.8%; Score 37.5; DB 1; Length 1196;  
Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
  
QY 1 RPKPQQWFW 9  
Db 837 QPNP-QWFW 844  
: | | | | |  
R P K P Q Q W F W 9  
  
RESULT 40  
ID DNBI\_HSV1F STANDARD; PRT; 1196 AA.  
AC P17469;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DE MAJOR DNA-BINDING PROTEIN (INFECTED CELL PROTEIN 8) (ICP 8 PROTEIN).  
GN DBP OR UL29 OR ICP8.  
OS Herpes simplex virus (type 1 / strain F).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10304;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88306231; PubMed=2841793;  
RA Hammerschmidt W., Conraths F., Mankertz J., Pauli G., Ludwig H.,  
RA Buhk H.-J.;  
RT "Conservation of a gene cluster including glycoprotein B in bovine  
RT herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).";  
RL Virology 165:388-405(1988).  
CC -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA

```
CC REPLICATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. IN THE ABSENCE OF DNA REPLICATION,
CC FOUND IN THE NUCLEAR FRAMEWORK-ASSOCIATED STRUCTURES
CC (PREREPLICATIVE SITES); AS VIRAL DNA REPLICATION PROCEEDS, IT
CC MIGRATES TO GLOBULAR INTRANUCLEAR STRUCTURES (REPLICATION
CC COMPARTMENTS).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
CC FAMILY.
CC PIR: D29242; DNBHF.
CC InterPro: IPR000635; Viral_DNA_bind.
CC Pfam: PF00747; viral_DNA_bp; 1.
CC DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
CC ZN_FING 499 512 C2HC-TYPE.
CC SEQUENCE 1196 AA; 128373 MW; BC872584DDB1C8E2 CRC64;

Query Match 52.8%; Score 37.5; DB 1; Length 1196;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPKPQQWF 9
Db 837 QPNP-QMFW 844

RESULT 41
DNBL_HSV1K
ID DNBL_HSV1K STANDARD; PRT; 1196 AA.
AC P17470;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MAJOR DNA-BINDING PROTEIN (INFECTED CELL PROTEIN 8) (ICP 8 PROTEIN).
GN DBP OR UL29 OR ICP8.
OS Herpes simplex virus (type 1 / strain KOS).
CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Simplexvirus.
CC NCBI_TaxID=10306;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=88179536; PubMed=2833010;
CC Gao M., Bouchey J., Curtin K., Knipe D.M.;
CC "Genetic identification of a portion of the herpes simplex virus ICP8
CC protein required for DNA-binding.";
CC Virology 163:319-329(1988).
CC -1- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
CC REPLICATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR. IN THE ABSENCE OF DNA REPLICATION,
CC FOUND IN THE NUCLEAR FRAMEWORK-ASSOCIATED STRUCTURES
CC (PREREPLICATIVE SITES); AS VIRAL DNA REPLICATION PROCEEDS, IT
CC MIGRATES TO GLOBULAR INTRANUCLEAR STRUCTURES (REPLICATION
CC COMPARTMENTS).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; M20165; AAA45793.1; -.
CC InterPro: IPR000635; Viral_DNA_bind.
CC Pfam: PF00747; viral_DNA_bp; 1.
CC DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
CC ZN_FING 499 512 C2HC-TYPE.
CC SEQUENCE 1196 AA; 128314 MW; OD0010A5DF0A4694 CRC64;

Query Match 52.8%; Score 37.5; DB 1; Length 1196;
MEDLINE=93228441; PubMed=8385914;
```

```
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPKPQQWF 9
Db 837 QPNP-QMFW 844

RESULT 42
DNBL_HSV2H
ID DNBL_HSV2H STANDARD; PRT; 1196 AA.
AC P89452;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MAJOR DNA-BINDING PROTEIN (INFECTED CELL PROTEIN 8) (ICP 8 PROTEIN).
GN DBP OR UL29 OR ICP8.
OS Herpes simplex virus (type 2 / strain HG52).
CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Simplexvirus.
CC NCBI_TaxID=10315;
CC [1]
CC SEQUENCE FROM N.A.
CC Dolan A.;
CC Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
CC REPLICATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC ENBL; 286099; CAB06754.1; -.
CC InterPro: IPR000635; Viral_DNA_bind.
CC Pfam: PF00747; viral_DNA_bp; 1.
CC DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
CC ZN_FING 499 512 C2HC-TYPE.
CC DOMAIN 1168 1196 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
CC SEQUENCE 1196 AA; 128423 MW; A19CA843280DD7F5 CRC64;

Query Match 52.8%; Score 37.5; DB 1; Length 1196;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPKPQQWF 9
Db 837 QPNP-QMFW 844

RESULT 43
DNBL_HSV2
ID DNBL_HSV2 STANDARD; PRT; 1197 AA.
AC P36384;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MAJOR DNA-BINDING PROTEIN (INFECTED CELL PROTEIN 8) (ICP 8 PROTEIN).
GN DBP OR UL29 OR ICP8.
OS Herpes simplex virus (type 2).
CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Simplexvirus.
CC NCBI_TaxID=10310;
CC [1]
CC SEQUENCE FROM N.A.
CC MEDLINE=93228441; PubMed=8385914;
```

RA Toh Y., Liu Y., Tanaka S., Mori R.;  
RT "Nucleotide sequence of the major DNA-binding protein gene of herpes  
simplex virus type 2 and a comparison with the type 1.";  
RL Arch. Virol. 129:183-196(1993).  
CC -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA  
REPLICATION.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN  
FAMILY.  
CC PIR: A48350; A48350.  
DR InterPro: IPR000635; Viral\_DNA\_bind.  
DR Pfam: PF00747; Viral\_DNA\_bp; 1.  
KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.  
FT ZN\_FING 499 512 C2HC-TYPE.  
FT DOMAIN 1169 1197 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
SQ SEQUENCE 1197 AA; 128412 MW; C1576B5B865BFB CRC64;

Query Match 52.8%; Score 37.5; DB 1; Length 1197;  
Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPKQOQWFW 9  
: : : : :  
Db 837 QPNP-QWFWL 844

RESULT 44  
DNBI\_VZVD  
ID DNBI\_VZVD STANDARD; PRT; 1204 AA.  
AC P09246;  
DT 01-MAR-1989 (Rel. 10, Created)  
DT 01-MAR-1989 (Rel. 10, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MAJOR DNA-BINDING PROTEIN.  
GN 29.  
OS Varicella-zoster virus (strain Dumas) (VZV).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=10338;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86306657; PubMed=3018124;  
RA Davison A.J., Scott J.E.;  
RT "The complete DNA sequence of varicella-zoster virus.";  
RL J. Gen. Virol. 67:1759-1816(1986).  
CC -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA  
REPLICATION.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN  
FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL: X04370; CAA27912.1; -.  
DR PIR: C27214; DNB229.  
DR InterPro: IPR000635; Viral\_DNA\_bind.  
DR Pfam: PF00747; Viral\_DNA\_bp; 1.  
KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.  
FT ZN\_FING 497 510 C2HC-TYPE.  
SQ SEQUENCE 1204 AA; 132139 MW; D2FEE65DC0CC674 CRC64;

Query Match 52.8%; Score 37.5; DB 1; Length 1204;  
Best Local Similarity 60.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PKPOQWFWLM 11  
: : : : :  
Db 836 PNP-QWFWTL 844

RESULT 45  
DNBI\_HSVB  
ID DNBI\_HSVB STANDARD; PRT; 1209 AA.  
AC P28932;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MAJOR DNA-BINDING PROTEIN.  
GN 31.  
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=31520;  
[1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92295566; PubMed=1318606;  
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;  
RT "The DNA sequence of equine herpesvirus-1.";  
RL Virology 189:304-316(1992).  
CC -!- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA  
REPLICATION.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).  
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN  
FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL: M86664; AAB02466.1; -.  
DR PIR: E36798; DNBE64.  
DR InterPro: IPR000635; Viral\_DNA\_bind.  
DR Pfam: PF00747; Viral\_DNA\_bp; 1.  
KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.  
FT ZN\_FING 503 516 C2HC-TYPE.  
SQ SEQUENCE 1209 AA; 129982 MW; 1A728FB04484FE95 CRC64;

Query Match 52.8%; Score 37.5; DB 1; Length 1209;  
Best Local Similarity 60.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PKPOQWFWLM 11  
: : : : :  
Db 844 PNP-QWFWTL 852

RESULT 46  
TKNA\_ONCMY  
ID TKNA\_ONCMY STANDARD; PRT; 11 AA.  
AC P28499;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE SUBSTANCE P.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
[1]  
RP SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=92298992; PubMed=1376687;

RA Jensen J., Conlon J.M.;  
RT "Substance-P-related and neurokinin-A-related peptides from the brain  
of the cod and trout.";  
RL Eur. J. Biochem. 206:659-664(1992).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR; S23307; S23307.  
DR PIR; S23308; S23308.  
DR InterPro; IPR003580; Protachykinin.  
DR InterPro; IPR002040; Tachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR SMART; SM00203; TK; 1.  
DR PROSITE; PS00267; TACHYKININ; 1.  
DR Tachykinin; Neuropeptide; Amidation; Neurotransmitter.  
KW Tachykinin; 11  
FT MOD\_RES 11  
SQ SEQUENCE 11 AA; 1358 MW; 214860DEC9D6D1F7 CRC64;  
  
Query Match 52.1%; Score 37; DB 1; Length 11;  
Best Local Similarity 54.5%; Pred. No. 5.8; Indels 2; Mismatches 3; Gaps 0;  
Matches 6; Conservative 3;  
  
QY 1 RPKPQOWFWLM 11  
: : : : :  
DB 1 KPRPHQFFGLM 11  
  
RESULT 47  
HOBB\_ECOLI  
ID HOBB\_ECOLI STANDARD; PRT; 175 AA.  
AC P36558;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE VERY HYPOTHETICAL HOBB PROTEIN.  
GN HOBB.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-K12 / W3110;  
RX MEDLINE=95009972; PubMed=7925311;  
RA Herrick J., Kern R., Guha S., Landoulsi A., Fayet O., Malki A.,  
RA Kohiyama M.;  
RT "Parental strand recognition of the DNA replication origin by the  
outer membrane in Escherichia coli.";  
RL EMBO J. 13:4695-4703(1994).  
RN [2]  
RP COMMENT ABOUT THIS PROTEIN.  
RA Balroch A.;  
RL Unpublished observations (NOV-1994).  
CC -1- FUNCTION: DNA-BINDING PROTEIN SPECIFIC OF THE E.COLI ORIGIN OF  
CC REPLICATION (ORIC).  
CC -1- CAUTION: THIS PROTEIN MAY NOT BE THE "REAL" HOBB BECAUSE IT IS ON  
CC THE OPPOSITE FRAME OF AN ORF (APHA) WHICH HAS BEEN SHOWN, BY  
CC MICROSEQUENCING, TO EXIST. FURTHERMORE THE REAL HOBB IS PROBABLY  
CC SEQA.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; Z26592; CAA81346.1; ALT\_SEQ.  
DR PIR; S37475; S37475.

KW Hypothetical protein; DNA-binding.  
SQ SEQUENCE 175 AA; 20876 MW; 3E8119754B0ADEB4 CRC64;  
  
Query Match 52.1%; Score 37; DB 1; Length 175;  
Best Local Similarity 71.4%; Pred. No. 70;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPKPOOW 7  
: : : : :  
DB 160 RPKPNEW 166  
  
RESULT 48  
HRB3\_XANCV  
ID HRB3\_XANCV STANDARD; PRT; 253 AA.  
AC P80152;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE HRPB3 PROTEIN PRECURSOR.  
GN HRPB3.  
OS Xanthomonas campestris (pv. vesicatoria).  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xanthomonas.  
OX NCBI\_TaxID=341;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-ISOLATE 75-3;  
RX MEDLINE=93113007; PubMed=1472717;  
RA Fenselau S., Balbo I., Bonas U.;  
RT "Determinants of pathogenicity in Xanthomonas campestris pv.  
vesicatoria are related to proteins involved in secretion in  
bacterial pathogens of animals.";  
RL Mol. Plant Microbe Interact. 5:390-396(1992).  
CC -1- FUNCTION: NECESSARY FOR BOTH BASIC PATHOGENICITY AND THE INDUCTION  
CC OF THE HYPERSENSITIVE RESPONSE IN RESISTANT PLANTS. COULD BE A  
CC PART OF A SPECIFIC TRANSPORT APPARATUS OR A SECRETICIN APPARATUS  
CC THAT IS REQUIRED FOR PATHOGENICITY. HRP PROTEINS MAY FORM A  
CC COMPLEX (TUNNEL/PORE) THAT ENABLES THE EXPORT OF MOLECULES SUCH AS  
CC VIRULENCE AND AVIRULENCE FACTORS.  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE OUTER MEMBRANE BY A LIPID  
CC ANCHOR (PROBABLE).  
CC -1- SIMILARITY: BELONGS TO THE YSCJ FAMILY OF LIPOPROTEINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U33548; AAB08458.1; -;  
DR EMBL; M99175; AAA27604.1; -;  
DR InterPro; IPR003282; SecIIOMPK.  
DR InterPro; IPR002920; YscJ\_Flip.  
DR Pfam; PF01514; YscJ\_Flip; 1.  
DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; 1.  
KW Transport; Protein transport; Outer membrane; Signal; Lipoprotein;  
KW Hypersensitive response.  
FT SIGNAL 1 18  
POTENTIAL.  
FT CHAIN 19 253  
HRPB3 PROTEIN.  
FT LIPID 19 19  
N-ACYL DIGLYCERIDE (POTENTIAL).  
SQ SEQUENCE 253 AA; 27263 MW; 35A80820C2D3555A CRC64;  
  
Query Match 52.1%; Score 37; DB 1; Length 253;  
Best Local Similarity 55.6%; Pred. No. 98;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PKPOOWFWL 10  
: : : : :  
: : : : :  
: : : : :



DR 204 PRPSPWPL 212

RESULT 49

FUT7\_HUMAN

AC Q11130; STANDARD; PRT; 342 AA.

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE ALPHA-1,3-FUCOSYLTRANSFERASE (EC 2.4.1.-) (GALACTOSIDE 3-L-FUCOSYLTRANSFERASE) (FUCOSYLTRANSFERASE 7) (FUCT-VII) (SELECTIN-LIGAND SYNTHASE).

GN FUT7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94266898; PubMed=8207002;

RA Natsuka S., Gersten K.M., Zenita K., Kannagi R., Lowe J.B.; "Molecular cloning of a cDNA encoding a novel human leukocyte alpha-1,3-fucosyltransferase capable of synthesizing the sialyl Lewis x determinant.";

RT J. Biol. Chem. 269:16789-16794(1994).

RL [2]

RN [3]

RP REVISIONS.

RX MEDLINE=94327669; PubMed=8051184;

RA Natsuka S., Gersten K.M., Zenita K., Kannagi R., Lowe J.B.; "Molecular cloning of a cDNA encoding a novel human leukocyte alpha-1,3-fucosyltransferase capable of synthesizing the sialyl Lewis x determinant.";

RT J. Biol. Chem. 269:20806-20806(1994).

RL [3]

RN [4]

RP SEQUENCE FROM N.A.

RA Hiraiwa N., Hiraiwa M., Kannagi R.; "The human selectin-ligand synthase (hFuc-T VII) gene structure and characterization of the promoter.";

RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: MAY CATALYZE ALPHA-1,3 GLYCOSIDIC LINKAGES INVOLVED IN THE EXPRESSION OF SIALYL LEWIS X ANTIGENS.

CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + ALPHA-2,3-NEU-N-ACETYL-1,4-BETA-D-GALACTOSYL-N-ACETYL-D-GLUCOSAMINYL-R = GDP + ALPHA-2,3-NEU-N-ACETYL-1,4-BETA-D-GALACTOSYL-(ALPHA-1,3-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINYL-R.

CC -!- PATHWAY: GLYCOSYLATION.

CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND FORM IN TRANS CISTERNAE OF GOLGI.

CC -!- TISSUE SPECIFICITY: LEUKOCYTIC/MYELOID LINEAGE CELLS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X78031; CAA54962.1; -

DR EMBL; U11282; AAA20468.1; -

DR EMBL; U08112; AAA56869.1; -

DR EMBL; AB012668; BAA32819.1; -

DR MIM; 602030; -

DR InterPro; IPR001503; Glyco\_transf\_10.

DR Pfam; PF00852; Glyco\_transf\_10; 1.

KW Transferase; Glycosyltransferase; Transmembrane; Glycoprotein;

KW Signal-anchor; Golgi stack.

FT DOMAIN 1 14 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 15 36 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (POTENTIAL).

FT DOMAIN 37 342 LUMENAL, CATALYTIC (POTENTIAL).

FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 161 162 GP -> A (IN REF. 1; AAA56869).

FT CONFLICT 304 305 RL -> SV (IN REF. 1; AAA56869).

SQ SEQUENCE 342 AA; 39238 MW; D31BFF90DD64DFAB CRC64;

Query Match 52.1%; Score 37; DB 1; Length 342;

Best Local Similarity 55.6%; Pred. No. 1.3e-02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPOQWFW 9

DB 110 RPRGQFWW 118

II: | | |

RESULT 50

YCGF\_ECOLI

ID YCGF\_ECOLI STANDARD; PRT; 403 AA.

AC P75990;

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE HYPOTHETICAL PROTEIN YCGF.

GN YCGF OR B1163.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;

RA "The complete genome sequence of Escherichia coli K-12.";

RT Science 277:1453-1474(1997).

RL [2]

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=K12.

RX MEDLINE=97061202; PubMed=8905232;

RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikenoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N., Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T.;

RA "A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 min region on the linkage map.";

RT DNA Res. 3:137-155(1996).

RL -!- SIMILARITY: CONTAINS 1 DUF2 DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AF000215; AAC74247.1; -

DR EMBL; D90750; BAA35998.1; -

DR EMBL; D90751; BAA36002.1; -

```
DR EcoGene; EGI3887; ycgF.
DR InterPro: IPR001633; DUF2.
DR Pfam: PF00563; DUF2; 1.
DR SMART: SM00052; DUF2; 1.
KW Hypothetical protein; Complete proteome.
FT DOMAIN 159 394 DUF2.
SQ SEQUENCE 403 AA; 45295 MW; 57B662BEC10957DA CRC64;

Query Match 52.1%; Score 37; DB 1; Length 403;
Best Local Similarity 57.1%; Pred. No. 1.5e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQOWFWL 10
   I::I
Db 367 PEENWNL 373

RESULT 51
CYB_MARPO
ID CYB_MARPO STANDARD; PRT; 404 AA.
AC P26852;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME B.
GN COB OR CVTB.
OS Marchantia polymorpha (Liverwort).
OC Mitochondrion.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;
OC Marchantiales; Marchantiaceae; Marchantia.
OX NCBI_TaxID=3197;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92114051; PubMed=1731062;
RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,
RA Akashi K., Kaneage T., Ogura Y., Kohchi T., Ohyama K.;
RT "Gene organization deduced from the complete sequence of liverwort
RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant
RT mitochondrial genome.";
RL J. Mol. Biol. 223:1-7(1992).

CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M68929; AAC09441.1; -.
CC PIR; S25953; S25953.
CC Mendel; 2054; MARPO; cob; 1.
CC InterPro: IPR000179; Cyt_b_b6.
CC Pfam; PF00032; cytochrome_b_c; 1.
CC Pfam; PF00033; cytochrome_b_n; 1.
CC PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
CC PROSITE; PS00193; CYTOCHROME_B_OO; 1.
CC Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
KW METAL 85 85 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 99 99 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 186 186 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 200 200 IRON 1 (HEME B566 AXIAL LIGAND).

SQ SEQUENCE 404 AA; 45188 MW; AFFCA920DD4783A2 CRC64;

Query Match 52.1%; Score 37; DB 1; Length 404;
Best Local Similarity 54.5%; Pred. No. 1.5e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11
   I::I::I
Db 322 RPIHQKFFWLL 332

RESULT 52
NEC3_MOUSE
ID NEC3_MOUSE STANDARD; PRT; 655 AA.
AC P29121;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NEUROENDOCRINE CONVERTASE 3 PRECURSOR (EC 3.4.21.-) (NEC 3) (PC4)
DE (PROHORMONE CONVERTASE 3) (KEX2-LIKE ENDOPROTEASE 3).
DE PCSK4 OR NEC3 OR NEC-3.
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92210552; PubMed=1372895;
RA Nakayama K., Kim W.S., Torii S., Hosaka M., Nakagawa T.,
RA Ikemizu J., Baba T., Murakami K.;
RT "Identification of the fourth member of the mammalian endoprotease
RT family homologous to the yeast Kex2 protease. Its testis-specific
RT expression.";
RL J. Biol. Chem. 267:5897-5900(1992).

CC -!- FUNCTION: INVOLVED IN THE PROCESSING OF HORMONE AND OTHER PROTEIN
CC PRECURSORS AT SITES COMPRISED OF PAIRS OF BASIC AMINO ACID
CC RESIDUES.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE
CC SUBTILASE FAMILY. FURIN SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D01093; BAA00877.1; -.
CC PIR; A42151; A42151.
CC HSP; Q99405; IMPT.
CC MEROPS; S08.074; -.
CC MGD; MGI:97514; Pcsk4.
CC InterPro: IPR002884; P_domain.
CC InterPro: IPR000209; Peptidase_S8.
CC Pfam; PF01483; P; 1.
CC Pfam; PF00082; Peptidase_S8; 1.
CC PRINTS; PR00723; SUBTILISIN.
CC ProDom; PD000717; P_domain; 1.
CC PROSITE; PS00136; SUBTILASE_ASP; 1.
CC PROSITE; PS00137; SUBTILASE_HIS; 1.
CC PROSITE; PS00138; SUBTILASE_SER; 1.
CC Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.
KW SIGNAL 1 26 BY SIMILARITY.
FT PROPEP 27 110 POTENTIAL.
FT CHAIN 111 655 NEUROENDOCRINE CONVERTASE 3.
FT DOMAIN 121 414 CATALYTIC.
FT ACT_SITE 155 155 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 196 196 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 370 370 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 472 472 N-LINKED (GLCNAC...) (POTENTIAL).
FT SSEQUENCE 655 AA; 73213 MW; 4E4E32CEDECB59 CRC64;
SQ SEQUENCE 655 AA; 73213 MW; 4E4E32CEDECB59 CRC64;
```

```
Query Match          52.1%; Score 37; DB 1; Length 655;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 QQWFWL 10
    |||:|
DB 625 QQWFWL 630

RESULT 53
TRA_BPMU          STANDARD;          PRT;          663 AA.
AC P07636; P06021;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TRANSPOSASE.
GN A OR 3.
OS Bacteriophage Mu.
OC Viruses; dsDNA viruses, no RNA stage; Tailed phages; Myoviridae.
OX NCBI_TaxID=10677;
RN [1]
RP MEDLINE=86067968; PubMed=2999776;
RA Harshey R.M., Getzoff E.D., Baldwin D.L., Miller J.L., Chaconas G.;
RT "Primary structure of phage mu transposase: homology to mu
  repressor.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:7676-7680(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA Priess H., Brauer B., Schmidt C., Kamp D.;
RT "Sequence of the left end of Mu.";
RL (In) Symonds N., Toussaint A., van de Putte P., Howe M.M. (eds.);
  Phage Mu, pp.277-296, Cold Spring Harbor Laboratory Press,
  New York (1987).
RN [3]
RP SEQUENCE FROM N.A.
RA Morgan G., Hatfull G., Hendrix R.;
RT "Genome of bacteriophage Mu and comparison with the Haemophilus
  influenzae Mu-like prophage Flumu.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 1-88 FROM N.A.
RX MEDLINE=83012203; PubMed=6214696;
RA Priess H., Kamp D., Kahmann R., Brauer B., Delius H.;
RT "Nucleotide sequence of the immunity region of bacteriophage Mu.";
RL Mol. Gen. Genet. 186:315-321(1982).
RN [5]
RP SEQUENCE OF 1-84 FROM N.A.
RX MEDLINE=83218562; PubMed=622246;
RA Toussaint A., Faalen M., Desmet L., Allet B.;
RT "The products of gene A of the related phages Mu and D108 differ in
  their specificities.";
RL Mol. Gen. Genet. 190:70-79(1983).
RN [6]
RP STRUCTURE BY NMR OF 1-76.
RX MEDLINE=95187707; PubMed=7881904;
RA Clubb R.T., Omichinski J.G., Savilahti H., Mizuuchi K.,
  Gronenborn A.M., Clore G.M.;
RT "A novel class of winged helix-turn-helix protein: the DNA-binding
  domain of Mu transposase.";
RL Structure 2:1041-1048(1994).
RN [7]
RP STRUCTURE BY NMR OF 76-174.
RX MEDLINE=98070329; PubMed=9405381;
RA Schumacher S., Clubb R.T., Cai M., Mizuuchi K., Clore G.M.,
  Gronenborn A.M.;
RT "Solution structure of the Mu end DNA-binding beta subdomain of
  phage Mu transposase: modular DNA recognition by two tethered
  domains.";
RL EMBO J. 16:7532-7541(1997).

[8]
RN RP STRUCTURE BY NMR OF 173-247.
RX MEDLINE=98035037; PubMed=93667742;
RA Clubb R.T., Schumacher S., Mizuuchi K., Gronenborn A.M., Clore G.M.;
RT "Solution structure of the I gamma subdomain of the Mu end
  DNA-binding domain of phage Mu transposase.";
RL J. Mol. Biol. 273:19-25(1997).
RN [9]
RN RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 248-574.
RX MEDLINE=95354202; PubMed=7628012;
RA Rice P., Mizuuchi K.;
RT "Structure of the bacteriophage Mu transposase core: a common
  structural motif for DNA transposition and retroviral integration.";
RL Cell 82:209-220(1995).
CC 1- FUNCTION: THIS TRANSPOSASE IS ESSENTIAL FOR INTEGRATION,
  REPLICATION-TRANSDUCTION, AND EXCISION OF MU DNA.
CC 2- MISCELLANEOUS: MU CAN TRANPOSE ITS DNA INTO MULTIPLE SITES IN
  MANY BACTERIAL GENOMES AND MEDIATE A VARIETY OF DNA
  REARRANGEMENTS. TRANSDUCTION REQUIRES BOTH TRANSPOSASE (ENCODED
  BY GENE A) AND TRANSDUCTION ENHANCER (ENCODED BY GENE B).
CC 3- MISCELLANEOUS: UNLIKE OTHER TRANSPOSONS MU HAS DISSIMILAR
  SEQUENCES AT ITS LEFT AND RIGHT ENDS. TRANSPOSASE APPARENTLY BINDS
  3 SPECIFIC BLOCKS OF SEQUENCES AT EACH END OF MU DNA.
CC 4- MISCELLANEOUS: THE A GENE IS REGULATED BY THE REPRESSOR C, WHICH
  BINDS TO AN OPERATOR SEQUENCE & TURNS OFF TRANSCRIPTION. REPRESSOR
  C CAN, AT HIGH CONCENTRATIONS, OCCUPY ALMOST THE EXACT SAME SITES
  ON MU ENDS AS THE TRANSPOSASE, AND TRANSPOSASE CAN BIND TO
  FRAGMENTS CONTAINING THE MU OPERATOR SEQUENCE.
CC 5- SIMILARITY: STRONG, TO H.INFLUENZAE H11478.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC
DR EMBL; M11195; AAA32369.1; -
DR EMBL; M64097; AAA32379.1; -
DR EMBL; AF083977; AAF01083.1; -
DR EMBL; V01464; CAA24713.1; -
DR EMBL; V00868; CAA24236.1; -
DR PIR; A24746; TOBPU.
DR PDB; 1TNS; 14-FEB-95.
DR PDB; 1TNT; 14-FEB-95.
DR PDB; 1BCM; 15-OCT-95.
DR PDB; 1BCO; 15-OCT-95.
DR PDB; 2EZH; 03-DEC-97.
DR PDB; 2EZI; 03-DEC-97.
DR PDB; 2E2K; 14-JAN-98.
DR PDB; 2E2L; 14-JAN-98.
DR InterPro; IPR003314; Mu_DNA_bind.
DR Pfam; PF02316; Mu_DNA_bind; 1.
KW Transposition; Transposable element; DNA-binding; DNA excision;
  DNA integration; DNA recombination; 3D-structure.
FT DNA_BIND 35 55 H-T-H MOTIF (POTENTIAL).
FT DNA_BIND 390 409 H-T-H MOTIF (POTENTIAL).
FT CONFLICT 66 66 G -> R (IN REF. 5).
FT CONFLICT 408 408 P -> S (IN REF. 2).
SQ SEQUENCE 663 AA; 75003 MW; B862CFDCBFC0B2E3 CRC64;

Query Match          52.1%; Score 37; DB 1; Length 663;
Best Local Similarity 57.1%; Pred. No. 2.3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPQQWF 9
    :|:|
DB 286 RPRTWF 292

RESULT 54
```

```

TBR1_MOUSE
ID AC Q64336; STANDARD; PRT; 681 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE T-BRAIN-1 PROTEIN (T-BOX BRAIN PROTEIN 1) (TBR-1) (TES-56).
GN TBR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Fetal brain;
RX MEDLINE=95344783; PubMed=7619531;
RA Bulfone A., Smiga S.M., Shimamura K., Peterson A., Puellès L.,
RA Rubenstein J.L.R.;
RT "T-brain-1: a homolog of Brachyury whose expression defines
RT molecularly distinct domains within the cerebral cortex.";
RL Neuron 15:63-78(1995).
CC -1- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
CC DEVELOPMENTAL PROCESSES. TBR1 IS REQUIRED FOR NORMAL BRAIN
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SPECIFIC LAMINA IN THE DEVELOPING
CC AND ADULT BRAIN.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AROUND DAY 10 OF EMBRYONIC
CC DEVELOPMENT IN THE PREPLATE, AT DAY 12.5, IN THE CORTICAL PLATE
CC AND INTERMEDIATE ZONE, AND FROM DAY 16.5 TO 18.5, IN A ROSTRO-
CC CAUDAL GRADIENT IN THE SUBPLATE. IN THE THALAMUS, EXPRESSION IS
CC FIRST OBSERVED AT POSTNATAL STAGE, P7, AND WEAK EXPRESSION
CC CONTINUES IN LATER POSTNATAL AND ADULT STAGES.
CC -1- SIMILARITY: CONTAINS A T-BOX DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL: U49251; AAA92011.1; --
HSSP: P24781; 1XBR.
MGI: 107404; Tbr1.
InterPro: IPR001699; T-box.
Pfam: PF00907; T-box; 1.
PRINTS: PR00937; TBOX.
SMART: SM00425; TBOX; 1.
PROSITE: PS01283; TBOX_1; 1.
PROSITE: PS01264; TBOX_2; 1.
PROSITE: PS0252; TBOX_3; 1.
Transcription regulation; DNA-binding; Nuclear protein.
DNA_BIND 213 393
DOMAIN 569 573 POLY-ALA.
SEQUENCE 681 AA; 73941 MW; 8732EF250EFID009 CRC64;

Query Match 52.1%; Score 37; DB 1; Length 681;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQWF 8
D 1 1 1 1 1
Db 479 PSQLWF 485

RESULT 55
TBR1_HUMAN
ID AC Q16650; STANDARD; PRT; 682 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE AMP DEAMINASE 1 (EC 3.5.4.6) (MYOADENYLATE DEAMINASE) (AMP DEAMINASE
DE ISOFORM M).
GN AMPD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hmo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

TBR1_MOUSE
ID AC Q64336; STANDARD; PRT; 681 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE T-BRAIN-1 PROTEIN (T-BOX BRAIN PROTEIN 1) (TBR-1) (TES-56).
GN TBR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Fetal brain;
RX MEDLINE=95344783; PubMed=7619531;
RA Bulfone A., Smiga S.M., Shimamura K., Peterson A., Puellès L.,
RA Rubenstein J.L.R.;
RT "T-brain-1: a homolog of Brachyury whose expression defines
RT molecularly distinct domains within the cerebral cortex.";
RL Neuron 15:63-78(1995).
CC -1- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
CC DEVELOPMENTAL PROCESSES. TBR1 IS REQUIRED FOR NORMAL BRAIN
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN SPECIFIC LAMINA IN THE DEVELOPING
CC AND ADULT BRAIN.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AROUND DAY 10 OF EMBRYONIC
CC DEVELOPMENT IN THE PREPLATE, AT DAY 12.5, IN THE CORTICAL PLATE
CC AND INTERMEDIATE ZONE, AND FROM DAY 16.5 TO 18.5, IN A ROSTRO-
CC CAUDAL GRADIENT IN THE SUBPLATE. IN THE THALAMUS, EXPRESSION IS
CC FIRST OBSERVED AT POSTNATAL STAGE, P7, AND WEAK EXPRESSION
CC CONTINUES IN LATER POSTNATAL AND ADULT STAGES.
CC -1- SIMILARITY: CONTAINS A T-BOX DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL: U49251; AAA92011.1; --
HSSP: P24781; 1XBR.
MGI: 107404; Tbr1.
InterPro: IPR001699; T-box.
Pfam: PF00907; T-box; 1.
PRINTS: PR00937; TBOX.
SMART: SM00425; TBOX; 1.
PROSITE: PS01283; TBOX_1; 1.
PROSITE: PS01264; TBOX_2; 1.
PROSITE: PS0252; TBOX_3; 1.
Transcription regulation; DNA-binding; Nuclear protein.
DNA_BIND 213 393
DOMAIN 569 573 POLY-ALA.
SEQUENCE 681 AA; 73941 MW; 8732EF250EFID009 CRC64;

Query Match 52.1%; Score 37; DB 1; Length 681;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQWF 8
D 1 1 1 1 1
Db 479 PSQLWF 485

RESULT 55
TBR1_HUMAN
ID AC Q16650; STANDARD; PRT; 682 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE AMP DEAMINASE 1 (EC 3.5.4.6) (MYOADENYLATE DEAMINASE) (AMP DEAMINASE
DE ISOFORM M).
GN AMPD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hmo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

TBR1_MOUSE
ID AC Q64336; STANDARD; PRT; 681 AA.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE T-BRAIN-1 PROTEIN (T-BOX BRAIN PROTEIN 1) (TBR-1) (TES-56).
GN TBR1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Fetal brain;
RX MEDLINE=95344783; PubMed=7619531;
RA Bulfone A., Smiga S.M., Shimamura K., Peterson A., Puellès L.,
RA Rubenstein J.L.R.;
RT "T-brain-1: a homolog of Brachyury whose expression defines
RT molecularly distinct domains within the cerebral cortex.";
RL Neuron 15:63-78(1995).
CC -1- FUNCTION: PROBABLE TRANSCRIPTIONAL REGULATOR INVOLVED IN
CC DEVELOPMENTAL PROCESSES. TBR1 IS REQUIRED FOR NORMAL BRAIN
CC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -1- TISSUE SPECIFICITY: BRAIN.
CC -1- SIMILARITY: CONTAINS A T-BOX DOMAIN.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to license@isb-sib.ch).
-----
EMBL: U49250; AAA92010.1; --
HSSP: P24781; 1XBR.
MIM: 604616; --
InterPro: IPR001699; T-box.
Pfam: PF00907; T-box; 1.
PRINTS: PR00937; TBOX.
SMART: SM00425; TBOX; 1.
PROSITE: PS01283; TBOX_1; 1.
PROSITE: PS01264; TBOX_2; 1.
PROSITE: PS0252; TBOX_3; 1.
Transcription regulation; DNA-binding; Nuclear protein.
DNA_BIND 213 393
DOMAIN 569 573 POLY-ALA.
SEQUENCE 682 AA; 74053 MW; ELC8D84206EFBBB5 CRC64;

Query Match 52.1%; Score 37; DB 1; Length 682;
Best Local Similarity 71.4%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQWF 8
D 1 1 1 1 1
Db 479 PSQLWF 485

RESULT 56
AMD1_HUMAN
ID AC P23109; STANDARD; PRT; 747 AA.
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE AMP DEAMINASE 1 (EC 3.5.4.6) (MYOADENYLATE DEAMINASE) (AMP DEAMINASE
DE ISOFORM M).
GN AMPD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Hmo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

RX MEDLINE=90264442; PubMed=2345176;  
RA Sabina R.L., Morisaki T., Clarke P., Eddy R., Shows T.B., Morton C.C.,  
RA Holmes E.W.;  
RT "Characterization of the human and rat myoadenylate deaminase  
RT genes";  
RL J. Biol. Chem. 265:9423-9433(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92131279; PubMed=1370861;  
RA Sabina R.L., Fishbein W.N., Pezeshkpour G., Clarke P.R., Holmes E.W.;  
RT "Molecular analysis of the myoadenylate deaminase deficiencies";  
RL Neurology 42:170-179(1992).  
CC -!- FUNCTION: AMP DEAMINASE PLAYS A CRITICAL ROLE IN ENERGY  
CC METABOLISM.  
CC -!- CATALYTIC ACTIVITY: AMP + H(2)O = IMP + NH(3).  
CC -!- PATHWAY: PURINE NUCLEOTIDE CYCLE.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- TISSUE SPECIFICITY: THREE ISOFORMS ARE PRESENT IN MAMMALS: AMP  
CC DEAMINASE 1 IS THE PREDOMINANT FORM IN SKELETAL MUSCLE; AMP  
CC DEAMINASE 2 PREDOMINATES IN SMOOTH MUSCLE, NON-MUSCLE TISSUE,  
CC EMBRYONIC MUSCLE AND UNDIFFERENTIATED MYOBLASTS; AMP DEAMINASE 3  
CC IS FOUND IN ERYTHROCYTES.  
CC -!- SIMILARITY: BELONGS TO THE ADENOSINE AND AMP DEAMINASES FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M37931; AAG24258.1; JOINED.  
DR EMBL; M37920; AAG24258.1; JOINED.  
DR EMBL; M37921; AAG24258.1; JOINED.  
DR EMBL; M37922; AAG24258.1; JOINED.  
DR EMBL; M37923; AAG24258.1; JOINED.  
DR EMBL; M37924; AAG24258.1; JOINED.  
DR EMBL; M37927; AAG24258.1; JOINED.  
DR EMBL; M37928; AAG24258.1; JOINED.  
DR EMBL; M37929; AAG24258.1; JOINED.  
DR EMBL; M37930; AAG24258.1; JOINED.  
DR EMBL; M60092; AAA57281.1; JOINED.  
DR MIM; 102770; -.  
DR InterPro: IPR001365; A.deaminase.  
DR Pfam: PF00962; A.deaminase; 1.  
DR PROSITE: PS00485; A.DEAMINASE; 1.  
KW Hydrolase; Nucleotide metabolism; Multigene family.  
FT ACT\_SITE 363 363 BY SIMILARITY.  
FT ACT\_SITE 573 573 BY SIMILARITY.  
FT ACT\_SITE 649 649 BY SIMILARITY.  
FT ACT\_SITE 650 650 BY SIMILARITY.  
SQ SEQUENCE 747 AA; 86489 MW; 1E15EBEE98B95763 CRC64;

Query Match 52.1%; Score 37; DB 1; Length 747;  
Best Local Similarity 83.3%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 PKPQW 7  
|||||  
Db 528 PKPQW 533

RESULT 57  
METE\_SOLSC  
ID METE\_SOLSC STANDARD; PRT; 764 AA.  
AC Q42662;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE 5-METHYLTETRAHYDROPTEROYLTRIGLUTAMATE--HOMOCYSTEINE METHYLTRANSFERASE  
DE (EC 2.1.1.14) (VITAMIN-B12-INDEPENDENT METHIONINE SYNTHASE ISOZYME)

DE (COBALAMIN-INDEPENDENT METHIONINE SYNTHASE ISOZYME).  
GN MET.  
OS Solenostemon scutellarioides (Coleus blumei).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Lamiales; Lamiaceae; Solenostemon.  
OX NCBI\_TaxID=4142;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RA Petersen M., Van der Straeten D., Bauw G.;  
RT "Full-length cDNA clone from Coleus blumei with high similarity to  
RT cobalamin-independent methionine synthase";  
RL (In) Plant Gene Register PGR95-049.  
CC -!- FUNCTION: CATALYZES THE TRANSFER OF A METHYL GROUP FROM 5-  
CC METHYLTETRAHYDROPTERATE TO HOMOCYSTEINE RESULTING IN METHIONINE  
CC FORMATION (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: 5-METHYLTETRAHYDROPTEROYL-L-GLUTAMATE + L-  
CC HOMOCYSTEINE = TETRAHYDROPTEROYL-L-GLUTAMATE + L-METHIONINE.  
CC -!- COFACTOR: ZINC (BY SIMILARITY).  
CC -!- PATHWAY: TERMINAL STEP IN THE DE NOVO BIOSYNTHESIS OF METHIONINE.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).  
CC -!- SIMILARITY: BELONGS TO THE VITAMIN-B12 INDEPENDENT METHIONINE  
CC SYNTHASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; Z49150; CAA89019.1; ALT INIT.  
DR Mendel; 11673; Colbl:1331:11673.  
DR InterPro: IPR002629; Methionine\_synth.  
DR Pfam; PF01717; Methionine\_synth; 2.  
KW Transferase; Methyltransferase; Methionine biosynthesis; Zinc.  
FT METAL 646 646 ZINC (BY SIMILARITY).  
FT METAL 648 648 ZINC (BY SIMILARITY).  
FT METAL 732 732 ZINC (BY SIMILARITY).  
SQ SEQUENCE 764 AA; 84589 MW; 43D65134C253602F CRC64;

Query Match 52.1%; Score 37; DB 1; Length 764;  
Best Local Similarity 66.7%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKPPQWFW 9  
|||||  
Db 533 RKPPMTVEW 541

RESULT 58  
NUTL\_MAGGR  
ID NUTL\_MAGGR STANDARD; PRT; 956 AA.  
AC Q01168;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE NITROGEN REGULATORY PROTEIN NUTL.  
GN NUTL.  
OS Magnaporthe grisea (Rice blast fungus) (Pyricularia grisea).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetes; Incertae sedis; Magnaportheaceae; Magnaporthe.  
OX NCBI\_TaxID=148305;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-GUYANA 11;  
RX MEDLINE=96335139; PubMed=8757395;  
RA Froeliger E.H., Carpenter B.E.;  
RT "NUT1, a major nitrogen regulatory gene in Magnaporthe grisea, is  
RT dispensable for pathogenicity".  
RL Mol. Gen. Genet. 251:647-656(1996).

CC -1- FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN; ACTIVATES EXPRESSION  
CC OF NITROGEN-REGULATED GENES.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U60290; AAB03415.1; -  
DR HSSP: P17429; 5GAT.  
DR InterPro: IPR000679; Znf\_GATA.  
DR Pfam: PF00320; GATA; 1.  
DR PRINTS; PR00619; GATAZNFINGER.  
DR SMART; SM00401; Znf\_GATA; 1.  
DR PROSITE; PS00344; GATA\_ZN\_FINGER\_1; 1.  
DR PROSITE; PS01114; GATA\_ZN\_FINGER\_2; 1.  
KW Transcription regulation; Activator; DNA-binding; Zinc-finger;  
KW Nuclear protein; Nitrate assimilation.  
FT ZN\_FING 663 687 GATA-TYPE.  
SQ SEQUENCE 956 AA; 100874 MW; 40ABDA5A07A7D7AB CRC64;  
  
Query Match 52.1%; Score 37; DB 1; Length 956;  
Best Local Similarity 71.4%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 4 PQOWFWL 10  
||:|:|  
DB 946 PQEWDL 952  
  
RESULT 59  
AREA\_GIBFU STANDARD; PRT; 971 AA.  
AC P78688;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 18-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE NITROGEN REGULATORY PROTEIN AREA.  
GN AREA.  
OS Gibberella fujikuroi (Bakanae and foot rot disease fungus) (Fusarium  
OS moniliforme).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Hypocreales; Nectriaceae; Gibberella.  
OX NCBI\_TaxID=5127;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=M567;  
RA Tudzynski B., Feng B., Marzluf G.A.;  
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: MAJOR NITROGEN REGULATORY PROTEIN. POSITIVELY ACTING  
CC REGULATORY GENE OF NITROGEN METABOLITE REPRESSION (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
CC -1- SIMILARITY: CONTAINS 1 GATA-TYPE ZINC FINGER.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Y11006; CAA71897.1; -  
DR HSSP: P17429; 5GAT.  
DR InterPro: IPR000679; Znf\_GATA.  
DR Pfam: PF00320; GATA; 1.  
DR PRINTS; PR00619; GATAZNFINGER.

DR SMART; SM00401; Znf\_GATA; 1.  
DR PROSITE; PS00344; GATA\_ZN\_FINGER\_1; 1.  
DR PROSITE; PS01114; GATA\_ZN\_FINGER\_2; 1.  
KW Transcription regulation; Activator; DNA-binding; Zinc-finger;  
KW Nuclear protein; Nitrate assimilation.  
FT ZN\_FING 694 718 GATA-TYPE.  
SQ SEQUENCE 971 AA; 103580 MW; 887DD882141C7453 CRC64;  
  
Query Match 52.1%; Score 37; DB 1; Length 971;  
Best Local Similarity 71.4%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 4 PQOWFWL 10  
||:|:|  
DB 961 PQEWL 967  
  
RESULT 60  
DOR\_DROME STANDARD; PRT; 1002 AA.  
AC Q24314;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DEEP ORANGE PROTEIN.  
GN DOR OR EG171E4.1.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97218037; PubMed=9065698;  
RA Shestopal S.A., Makulin I.V., Belyaeva E.S., Ashburner N.;  
RT "Molecular characterization of the deep orange (dor) gene of  
RT Drosophila melanogaster."  
RL Mol. Gen. Genet. 253:642-648(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OREGON-R;  
RX MEDLINE=20196011; PubMed=10731137;  
RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,  
RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,  
RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borkova D.,  
RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,  
RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablo B.,  
RA Modolell J., Peter A., Schoettler P., Werner M., Mourikoti F.,  
RA Beinhart N., Dowe G., Schaefer U., Jaeckle H., Bucheton A.,  
RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,  
RA McMillan P.J., Salles C., Tait E.A., Valenti P., Saunders R.D.C.,  
RA Glover D.M.;  
RT "From sequence to chromosome: the tip of the X chromosome of D.  
RT melanogaster."  
RL Science 287:2220-2222(2000).  
CC -1- SIMILARITY: SOME, TO YEAST PEP3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X86683; CAA60382.1; -  
DR EMBL: AL021726; CAA16809.1; -  
DR FlyBase; FBgn0000482; dor.  
DR InterPro: IPR000547; Clathrin\_repeat.  
DR InterPro: IPR001841; Znf\_ring.  
DR SMART; SM0299; CLH; 1.  
DR SMART; SM00184; RING; 1.

KW Zinc-finger; Transmembrane.  
FT ZN\_FING 885 910 C3H2C-TYPE.  
FT TRANSMEM 971 991 POTENTIAL.  
FT CONFLICT 169 169 A -> P (IN REF. 1).  
FT CONFLICT 581 581 Q -> H (IN REF. 1).  
FT CONFLICT 865 865 A -> V (IN REF. 1).  
SQ SEQUENCE 1002 AA; 115305 MW; D59690A0FC95182F CRC64;

Query Match 52.1%; Score 37; DB 1; Length 1002;  
Best Local Similarity 71.4%; Pred. No. 3.4e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQFWL 10  
I:|I|I|  
Db 301 PKQAWL 307

RESULT 61  
DPOG\_HUMAN  
ID DPOG\_HUMAN STANDARD; PRT; 1239 AA.  
AC P54098; O92515;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE DNA POLYMERASE GAMMA (EC 2.7.7.7) (MITOCHONDRIAL DNA POLYMERASE  
DE CATALYTIC SUBUNIT).  
GN POLG OR MDPI.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97038687; PubMed=8884268;  
RA Ropp P.A., Copeland W.C.;  
RT "Cloning and characterization of the human mitochondrial DNA  
RT polymerase, DNA polymerase gamma.";  
RL Genomics 36:449-458(1996).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97186710; PubMed=9034326;  
RA Lecrenier N.L., van der Bruggen P., Foury F.;  
RT "Mitochondrial DNA polymerases from yeast to man: a new family of  
RT polymerases.";  
RL Gene 185:147-152(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Watanabe T.K., Shimizu F., Nishino N., Fujiwara T., Kanemoto N.,  
RA Suzuki M., Nakamura Y., Hirai Y., Maekawa H., Takahashi E.;  
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: INVOLVED IN THE REPLICATION OF MITOCHONDRIAL DNA.  
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE =  
CC N PYROPHOSPHATE + DNA(N).  
CC -!- COFACTOR: MAGNESIUM.  
CC -!- SUBUNIT: HOMOTETRAMER.  
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.  
CC -!- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:  
CC ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR  
CC DIFFERENT REACTIONS OF DNA SYNTHESIS.  
CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U60325; AAC50712.1; -.  
DR EMBL; X98093; CAA66719.1; -.

DR EMBL; D84103; BAA12223.1; -.  
DR MIM; 174763; -.  
DR InterPro; IPR002297; DNA\_polG.  
DR InterPro; IPR001098; DNA\_pol\_A.  
DR Pfam; PF00476; DNA\_pol\_A; 1.  
DR PRINTS; PR00867; DNApolG.  
DR SMART; SM00482; POLAC; 1.  
DR PROSITE; PS00447; DNA\_POLYMERASE\_A; 1.  
KW Transferase; DNA-directed DNA polymerase; DNA replication;  
KW DNA-binding; Mitochondrion; Magnesium.  
FT DOMAIN 43 60 POLY-GLN.  
FT DOMAIN 535 538 POLY-GLU.  
FT CONFLICT 55 55 Q -> QQQ (IN REF. 3).  
SQ SEQUENCE 1239 AA; 139562 MW; 2D9ECCD75AD6E01E CRC64;  
  
Query Match 52.1%; Score 37; DB 1; Length 1239;  
Best Local Similarity 62.5%; Pred. No. 4.1e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PKPOQFW 9  
| | | | |  
Db 164 PKPPANAW 171  
  
RESULT 62  
ATS9\_HUMAN  
ID ATS9\_HUMAN STANDARD; PRT; 1629 AA.  
AC Q9P2N4; Q9NR29;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ADAM-TS 9 PRECURSOR (EC 3.4.24.-) (A DISINTEGRIN AND METALLOPROTEINASE  
DE WITH THROMBOSPONDIN MOTIFS 9) (ADAMTS-9) (ADAM-TS9).  
GN ADAMTS9 OR KIAA1312.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
RC TISSUE=Fetal;  
RX MEDLINE=20396138; PubMed=10936055;  
RA Clark M.E., Kelnner G.S., Turbeville L.A., Boyer A., Arden K.A.,  
RA Maki R.A.;  
RT "ADAMTS 9, a novel member of the ADAM-TS/Metallospodin gene  
RT family.";  
RL Genomics 67:343-350(2000).  
RN [2]  
RP SEQUENCE OF 159-1629 FROM N.A. (LONG ISOFORM).  
RC TISSUE=Brain;  
RX MEDLINE=20181126; PubMed=10718198;  
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XVI.  
RT The complete sequences of 150 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 7:65-73(2000).  
CC -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR  
CC MATRIX (BY SIMILARITY).  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ALL FETAL TISSUES.  
CC EXPRESSED SLIGHTLY IN ADULT OVARY, PANCREAS, HEART, KIDNEY, LUNG,  
CC PLACENTA. ALSO DETECTED IN SPINAL CORD AND BRAIN. NOT DETECTED IN  
CC COLON, SMALL INTESTINE, TESTIS, LIVER, SKELETAL MUSCLE, SPLEEN OR  
CC THYMUS.  
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT  
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY  
CC SIMILARITY).  
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B (ZINC

CC METALLOPROTEASE); ALSO KNOWN AS THE REPOLYSIN SUBFAMILY.  
CC -1- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 11 TSP TYPE-1 DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF261918; AAF89106.1; -  
CC EMBL; AB037733; BAA92550.1; -  
CC MIM; 605421; -  
CC InterPro; IPR001590; Repolysin.  
CC InterPro; IPR000884; TSPL.  
CC InterPro; IPR000130; Zn\_MTPptdse.  
CC Pfam; PF01421; Repolysin; 1.  
CC Pfam; PF00090; tsp\_1; 11.  
CC SMART; SM00209; TSPL; 13.  
CC PROSITE; PS02115; ADAM\_MPRO; 1.  
CC PROSITE; PS00427; DISINTEGRINS; FALSE\_NEG.  
CC PROSITE; PS00092; TSPL; 9.  
CC PROSITE; PS00142; ZINC\_PROTEASE; 1.  
CC Hydrolase; Metalloprotease; Zinc; Signal; Glycoprotein; Zymogen;  
KW Repeat; Extracellular matrix; Alternative splicing.  
FT SIGNAL 1 18 POTENTIAL.  
FT PROPEP 19 287 BY SIMILARITY.  
FT CHAIN 288 1629 ADAM-TS 9.  
FT DOMAIN 509 587 DISINTEGRIN-LIKE.  
FT DOMAIN 589 642 TSP TYPE 1 1.  
FT DOMAIN 645 752 CYS-RICH.  
FT DOMAIN 753 880 SPACER.  
FT DOMAIN 999 1053 TSP TYPE 1 2.  
FT DOMAIN 1056 1108 TSP TYPE 1 3.  
FT DOMAIN 1111 1156 TSP TYPE 1 4.  
FT DOMAIN 1184 1239 TSP TYPE 1 5.  
FT DOMAIN 1240 1295 TSP TYPE 1 6.  
FT DOMAIN 1332 1383 TSP TYPE 1 7.  
FT DOMAIN 1385 1439 TSP TYPE 1 8.  
FT DOMAIN 1445 1498 TSP TYPE 1 9.  
FT DOMAIN 1501 1554 TSP TYPE 1 10.  
FT DOMAIN 1562 1612 TSP TYPE 1 11.  
FT DOMAIN 88 96 POLY-SER.  
FT SITE 223 223 CYSTEINE SWITCH (POTENTIAL).  
FT METAL 434 434 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 435 435 BY SIMILARITY.  
FT METAL 438 438 ZINC (CATALYTIC) (BY SIMILARITY).  
FT METAL 444 444 ZINC (CATALYTIC) (BY SIMILARITY).  
FT CARBOHYD 112 112 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 135 135 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 749 749 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 840 840 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1213 1213 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 1267 1267 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT VARSPLIC 1064 1072 CLVTCCKGH -> VRWEGCYFP (IN  
FT VARSPLIC 1073 1072 SHORT ISOFORM).  
FT CONFLICT 367 367 MISSING (IN SHORT ISOFORM).  
FT SEQUENCE 1629 AA; 182649 MW; C1C4CEFF58B8941F CRC64;  
Query Match 52.1%; Score 37; DB 1; Length 1629;  
Best Local Similarity 71.4%; Pred. NO. 5.3e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 KPQQWF 9  
| | | | |  
DB 875 KPQQFW 881

RESULT 63  
CPF6\_RAT STANDARD; PRT; 537 AA.  
AC P51871.  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE CYTOCHROME P450 4F6 (EC 1.14.14.1) (CYP1B6).  
GN CYP4F6.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Brain;  
RA MEDLINE=96125358; PubMed=8554568;  
RX Kawashima H., Strobel H.W.;  
RT "cDNA cloning of three new forms of rat brain cytochrome P450  
RT belonging to the CYP4F subfamily";  
RL Biochem. Biophys. Res. Commun. 217:1137-1144(1995).  
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN LIVER AND KIDNEY. LOWER  
CC EXPRESSION IN BRAIN.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; U39208; AAC52360.1; -  
CC InterPro; IPR001128; Cyt\_P450.  
CC Pfam; PF00067; P450; 1.  
CC PRINTS; PR00385; P450.  
CC PRINTS; PR00464; EP450II.  
CC PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT BINDING 468 468 HEME (BY SIMILARITY).  
FT SEQUENCE 537 AA; 61541 MW; 4D96D761A2BEA7E9 CRC64;  
Query Match 51.4%; Score 36.5; DB 1; Length 537;  
Best Local Similarity 75.0%; Pred. NO. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
QY 2 KPQQWF 9  
| | | | |  
DB 55 KPQ-SFW 61  
RESULT 64  
NOX1\_HUMAN STANDARD; PRT; 564 AA.  
AC Q9Y5S8; O95691;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE NADPH OXIDASE HOMOLOG 1 (NOX-1) (NADH/NADPH MITOGENIC  
DE OXIDASE SUBUNIT P65-MOX) (MITOGENIC OXIDASE 1) (MOX1).  
GN NOX1 OR NOH1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM NOH-1L).



RC TISSUE-Colon epithelium;  
RX MEDLINE=99413719; PubMed=10485709;  
RA Suh Y.-A., Arnold R.S., Lassegue B., Shi J., Xu X., Sorescu D.,  
RA Chung A.B., Griendling K.K., Lambeth J.D.;  
RT "Cell transformation by the superoxide-generating oxidase Mox1.";  
RL Nature 401:79-82(1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS NOH-1L; NOH-1LV AND NOH-1S).  
RX MEDLINE=20082959; PubMed=10615049;  
RA Banfi B., Maturana A., Jaconi S., Arnaudeau S., Laforge T., Sinha B.,  
RA Ligeti E., Denaurex N., Krause K.-H.;  
RT "A mammalian H+ channel, generated through alternative splicing of the  
RT NADPH oxidase homolog NOH-1.";  
RL Science 287:138-142(2000).  
RN [3]  
RP SEQUENCE OF 16-564 FROM N.A. (ISOFORM NOH-1L).  
RA Lloyd D.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: NOH-1S IS A VOLTAGE-GATED PROTON CHANNEL THAT MEDIATES  
CC THE H+ CURRENTS OF RESTING PHAGOCYTES AND OTHER TISSUES. IT  
CC PARTICIPATES IN THE REGULATION OF CELLULAR PH AND IS BLOCKED BY  
CC ZINC. NOH-1L IS A PYRIDINE NUCLEOTIDE-DEPENDENT OXIDOREDUCTASE  
CC THAT GENERATES SUPEROXIDE AND MIGHT CONDUCT H+ IONS AS PART OF ITS  
CC ELECTRON TRANSPORT MECHANISM, WHEREAS NOH-1S DOES NOT CONTAIN AN  
CC ELECTRON TRANSPORT CHAIN.  
CC -!- COFACTOR: NADP AND FAD (POTENTIAL).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; NOH-1L (SHOWN HERE), NOH-1S AND  
CC NOH-1LV; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- TISSUE SPECIFICITY: NOH-1L IS DETECTED IN COLON, UTERUS, PROSTATE,  
CC AND COLON CARCINOMA, BUT NOT IN PERIPHERAL BLOOD LEUKOCYTES. NOH-  
CC 1S IS DETECTED ONLY IN COLON AND COLON CARCINOMA CELLS.  
CC -!- SIMILARITY: BELONGS TO THE FRE / CYBB FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; AF127763; AAD38133.1; -  
DR EMBL; AF166326; AAF23232.1; -  
DR EMBL; AF166327; AAF23233.1; -  
DR EMBL; AF166328; AAF23234.1; -  
DR EMBL; Z83619; CAB06073.1; ALT\_SEQ.  
DR MIM; 300225; -  
DR InterPro; IPR002916; Ferric\_reduct.  
DR InterPro; IPR000778; GP91PHOX.  
DR Pfam; PF01794; Ferric\_reduct; 1.  
DR PRINTS; PR00466; GP91PHOX.  
KW Oxidoreductase; NADP; Electron transport; Transmembrane; FAD; Heme;  
KW Glycoprotein; Voltage-gated channel; Ionic channel;  
KW Alternative splicing.  
FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 10 30 POTENTIAL.  
FT DOMAIN 31 44 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 45 72 POTENTIAL.  
FT DOMAIN 73 102 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 103 123 POTENTIAL.  
FT DOMAIN 124 168 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 169 189 POTENTIAL.  
FT DOMAIN 190 206 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 207 227 POTENTIAL.  
FT DOMAIN 228 396 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 397 417 POTENTIAL.  
FT DOMAIN 418 564 CYTOPLASMIC (POTENTIAL).  
FT NP\_BIND 338 344 FAD (POTENTIAL).  
FT BINDING 101 101 HEME (PROBABLE).  
FT BINDING 115 115 HEME (PROBABLE).  
FT BINDING 209 209 HEME (PROBABLE).  
FT BINDING 221 221 HEME (PROBABLE).

FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARSPLIC 159 190 QSRNTVEYVTFSTIAGLTGVIMTIALILMTV -> HPHIT  
FT PTVMFTVTFDMLSSVNSNLLFLLIK (IN ISOFORM  
FT NOH-1S).  
FT VARSPLIC 191 564 MISSING (IN ISOFORM NOH-1S).  
FT VARSPLIC 433 481 MISSING (IN ISOFORM NOH-1LV).  
FT CONFLICT 173 173 I -> V (IN REF. 2).  
SQ SEQUENCE 564 AA; 64870 MW; C3BE290FAE6DBC9A CRC64;  
  
Query Match 51.4%; Score 36.5; DB 1; Length 564;  
Best Local Similarity 37.5%; Pred. No. 2.4e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 5; Gaps 1;  
  
QY 1 RPK-----PQOQFWLW 11  
||| | : | : :  
DB 259 RPKFEGHPPEMKWIL 274  
  
RESULT 65  
Y102\_MYCLE  
ID Y102\_MYCLE STANDARD; PRT; 659 AA.  
AC P53525;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL 71.2 KDA PROTEIN ML1998.  
GN ML1998 OR O659.  
OS Mycobacterium leprae.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96059637; PubMed=7476188;  
RA Fsihi H., Cole S.T.;  
RT "The Mycobacterium leprae genome: systematic sequence analysis  
RT identifies key catabolic enzymes, ATP-dependent transport systems and  
RT a novel *poa* locus associated with genomic variability.";  
RL Mol. Microbiol. 16:909-919(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TN;  
RX MEDLINE=21128732; PubMed=11234002;  
RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,  
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,  
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,  
RA Murphy L., Oliver K., Quail M.A., Rajandream M.-A., Rutherford K.M.,  
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
RA Barrell B.G.;  
RT "Massive gene decay in the leprosy bacillus.";  
RL Nature 409:1007-1011(2001).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -!- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV0102.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL; Z46257; CAA86362.1; -  
DR EMBL; AL583924; CAC30953.1; -  
DR Leproma; ML1998; -  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 24 44 POTENTIAL.  
FT TRANSMEM 71 91 POTENTIAL.

```
FT TRANSEM 115 135 POTENTIAL.
FT TRANSEM 157 177 POTENTIAL.
FT TRANSEM 183 203 POTENTIAL.
FT TRANSEM 214 234 POTENTIAL.
FT TRANSEM 242 262 POTENTIAL.
FT TRANSEM 279 299 POTENTIAL.
FT TRANSEM 311 331 POTENTIAL.
FT TRANSEM 365 385 POTENTIAL.
FT TRANSEM 393 413 POTENTIAL.
FT TRANSEM 433 453 POTENTIAL.
FT TRANSEM 490 510 POTENTIAL.
FT TRANSEM 517 537 POTENTIAL.
FT TRANSEM 550 570 POTENTIAL.
FT TRANSEM 596 616 POTENTIAL.
SQ SEQUENCE 659 AA; 71246 MW; 2C1D9C6BF720E39 CRC64;

Query Match 51.4%; Score 36.5; DB 1; Length 659;
Best Local Similarity 50.0%; Pred. No. 2.7e+02;
Matches 6; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 RPKPOQWF-WLM 11
   :|:|:|
Db 464 QPGPWEWTWLM 475

RESULT 66
DNBL_HSVB2
ID DNBL_HSVB2 STANDARD; PRT; 1186 AA.
AC P12639;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MAJOR DNA-BINDING PROTEIN.
GN DBP OR UL29.
OS Bovine herpesvirus type 2 (strain BWV) (Bovine mammillitis virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10296;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88306231; PubMed=2841793;
RA Hammerschmidt W., Conraths F., Mankertz J., Pauli G., Ludwig H.,
RA Buhk H.-J.;
RT "Conservation of a gene cluster including glycoprotein B in bovine
RT herpesvirus type 2 (BHV-2) and herpes simplex virus type 1 (HSV-1).";
RL Virology 165:388-405(1988).
RN [2]
RP SEQUENCE OF 1058-1186 FROM N.A.
RX MEDLINE=88306232; PubMed=2457278;
RA Hammerschmidt W., Conraths F., Mankertz J., Buhk H.-J., Pauli G.,
RA Ludwig H.;
RT "Common epitopes of glycoprotein B map within the major DNA-binding
RT proteins of bovine herpesvirus type 2 (BHV-2) and herpes simplex
RT virus type 1 (HSV-1).";
RL Virology 165:406-418(1988).
CC -1- FUNCTION: SINGLE-STRAND DNA-BINDING PROTEIN REQUIRED FOR DNA
CC REPLICATION.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES DNA-BINDING PROTEIN
CC FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M21630; AAA46051.1;
PIR; A29242; DNBEBG.
HSSP; P25816; ICQA.
```

```
DR InterPro: IPR000635; Viral_DNA_bind.
DR Pfam: PF00747; Viral_DNA_bp; 1.
KW DNA-binding; DNA replication; Zinc-finger; Nuclear protein.
FT ZN_FING 495 508 C2HC-TYPE.
SQ SEQUENCE 1186 AA; 127286 MW; A586ECC1479FBD2C CRC64;

Query Match 51.4%; Score 36.5; DB 1; Length 1186;
Best Local Similarity 75.0%; Pred. No. 4.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 RPKPOQWF 9
   :|:|:|
Db 834 PNP-QMFW 840

RESULT 67
TKNA_GADMO
ID TKNA_GADMO STANDARD; PRT; 11 AA.
AC P28498;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE SUBSTANCE P.
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadoidae; Gadidae;
OC Gadus.
OX NCBI_TaxID=8049;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=92298992; PubMed=1376687;
RA Jensen J., Conlon J.M.;
RT "Substance P-related and neurokinin-A-related peptides from the brain
RT of the cod and trout.";
RL Eur. J. Biochem. 206:659-664(1992).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
DR InterPro: IPR003580; Protachykinin.
DR InterPro: IPR002040; Tachykinin.
DR Pfam: PF02202; Tachykinin; 1.
DR SMART: SM00203; TK; 1.
DR PROSITE: PS00267; TACHYKININ; 1.
KW Tachykinin; Neuropeptide; Amidation; Neurotransmitter.
FT MOD_RES 11 11 AMIDATION (BY SIMILARITY).
SQ SEQUENCE 11 AA; 1315 MW; 214860D759D6C6C7 CRC64;

Query Match 50.7%; Score 36; DB 1; Length 11;
Best Local Similarity 54.5%; Pred. No. 8;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPOQWF 11
   :|:|:|
Db 1 KRPQQFGLM 11

RESULT 68
ATP8_DICLA
ID ATP8_DICLA STANDARD; PRT; 55 AA.
AC Q36362;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP SYNTHASE PROTEIN 8 (EC 3.6.1.34) (ATPASE SUBUNIT 8) (A6L).
GN ATP8 OR ATP8.
OS Dicertrarchus labrax (European sea bass).
OG Mitochondrion.
```



```
RESULT 71
MSBB_HAEIN STANDARD; PRT; 318 AA.
AC P44567;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LIPID A BIOSYNTHESIS (KDO)2-(LAUROYL)-LIPID IVA ACYLTRANSFERASE
DE (EC 2.3.1.-)
GN MSBB OR HI0199.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann A.R., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Karvagh A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus
RT influenzae Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: TRANSFERS MYRISTATE OR LAURATE, ACTIVATED ON ACP, TO
CC (KDO)2-(LAUROYL)-LIPID IVA (BY SIMILARITY).
CC -1- PATHWAY: LIPOPOLYSACCHARIDE CORE BIOSYNTHESIS (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE HTRB/MSBB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32705; AAC21868.1; --
DR TIGR; HI0199; -.
KW Lipopolysaccharide biosynthesis; Transferase; Acyltransferase;
KW Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 27 47 POTENTIAL.
FT TRANSMEM 91 111 POTENTIAL.
FT TRANSMEM 138 158 POTENTIAL.
SQ SEQUENCE 318 AA; 36882 MW; DE59952D78719445 CRC64;

Query Match 50.7%; Score 36; DB 1; Length 318;
Best Local Similarity 40.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQWFWM 11
| | | |
Db 295 PAPEQYVWIL 304

RESULT 72
Y355_SYNY3 STANDARD; PRT; 330 AA.
AC P74436;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)

Query Match 50.7%; Score 36; DB 1; Length 318;
Best Local Similarity 40.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQWFWM 11
| | | |
Db 295 PAPEQYVWIL 304
```

```
DE HYPOTHETICAL 36.1 KDA PROTEIN SLL0355.
GN SLL0355.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugiura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. PCC6803. II. Sequence determination of the entire
RT genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -1- SIMILARITY: BELONGS TO THE EMAA TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90915; BAA18537.1; --
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 2.
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 15 35 POTENTIAL.
FT TRANSMEM 41 61 POTENTIAL.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 125 145 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 201 221 POTENTIAL.
FT TRANSMEM 238 258 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 286 306 POTENTIAL.
SQ SEQUENCE 330 AA; 36092 MW; AB44A99D8B8B53DAC CRC64;

Query Match 50.7%; Score 36; DB 1; Length 330;
Best Local Similarity 71.4%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRPQOW 7
| | | |
Db 64 RRPQOW 70

RESULT 73
WNSB_HUMAN STANDARD; PRT; 359 AA.
AC Q9H1J7;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE WNT-5B PROTEIN PRECURSOR.
GN WNT5B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Testa T.T., Mossakowski D.E., Carter P.S., Hu E., Zhu Y.,
RA Kelsell D.P., Murdock P.R., Herrity N.C., Lewis C.J., Cross D.A.,
RA Culbert A.A., Reith A.D., Barnes M.R.;
RT "Molecular cloning and characterization of six novel human WNT
```



```
CC EMBL; X73510; CAA51916.1; -
DR EMBL; X76190; CAA53784.1; -
DR PIR; S34173; S34173; Wnt1.
DR InterPro; IPR000970; Wnt1.
DR Pfam; PF00110; wnt; 1.
DR SMART; SM00097; WNT1; 1.
DR PROSITE; PS00246; WNT1; 1.
KW Developmental protein; Glycoprotein; Signal; Extracellular matrix.
FT SIGNAL 1 16
FT CHAIN 17 360
FT CARBOHYD 94 94
FT CARBOHYD 100 100
FT CARBOHYD 292 292
FT CARBOHYD 306 306
FT CARBOHYD 306 306
FT CONFLICT 15 15
FT SEQUENCE 360 AA; 40714 MW; 93CBD15D7A92779E CRC64;

Query Match 50.7%; Score 36; DB 1; Length 360;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 RPK--PQWFW 9
III I: I
DB 149 RPKDPRDLWLW 159

RESULT 76
FUT3_BOVIN STANDARD; PRT; 365 AA.
AC Q11126;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE GALACTOSIDE 3(4)-L-FUCOSYLTRANSFERASE (EC 2.4.1.65) (BLOOD GROUP LEWIS
DE ALPHA-4-FUCOSYLTRANSFERASE) (LEWIS FT) (FUCOSYLTRANSFERASE 3) (FUCT-
DE III) (FUTB).
GN FUT3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP MEDLINE=97236840; PubMed=9079712;
RA Oulmoudene A., Wierinckx A., Petit J.-M., Costache M., Palcic M.M.,
RA Mollicon R., Oriol R., Julien R.;
RT "Molecular cloning and expression of a bovine alpha(1,3)-
RT fucosyltransferase gene homologous to a putative ancestor gene of the
RT human FUT3-FUT5-FUT6 cluster."
RT J. Biol. Chem. 272:8764-8773(1997).
CC -1- FUNCTION: MAY CATALYSE ALPHA-1,3 AND ALPHA-1,4 GLYCOSIDIC LINKAGES
CC INVOLVED IN THE EXPRESSION OF SIALYL LEWIS X AND LEWIS X/SSA-1
CC ANTIGENS. IT MAY BE INVOLVED IN BLOOD GROUP LEWIS DETERMINATION.
CC -1- CATALYTIC ACTIVITY: GDP-L-FUCOSE + 1,3-BETA-D-GALACTOSYL-
CC N-ACETYL-D-GLUCOSAMINYL-R = GDP + 1,3-BETA-D-GALACTOSYL-
CC (ALPHA-1,4-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINYL-R.
CC -1- PATHWAY: GLYCOSYLATION
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -1- TISSUE SPECIFICITY: LIVER, KIDNEY, LUNG AND BRAIN.
CC -1- MISCELLANEOUS: ALSO ACTS ON THE CORRESPONDING 1,4-GALACTOSYL
CC DERIVATIVE, FORMING 1,3-L-FUCOSYL LINKS.
CC -1- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN
CC GLYCOSYLTRANSFERASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X87810; CAA61079.1; -
DR InterPro; IPR001503; Glyco_transf_10.
DR Pfam; PF00852; Glyco_transf_10; 1.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 14
FT TRANSMEM 15 34
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 100 100
FT CARBOHYD 158 158
FT CARBOHYD 189 189
FT CARBOHYD 189 189
FT N-LINKED (GLCNAC. . .) (PROBABLE).
FT N-LINKED (GLCNAC. . .) (PROBABLE).
FT SEQUENCE 365 AA; 42654 MW; 18715A361B0025D3 CRC64;

Query Match 50.7%; Score 36; DB 1; Length 365;
Best Local Similarity 55.6%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQWFW 9
III I: I
DB 130 RPPQRWVW 138

RESULT 77
WNT5A_HUMAN STANDARD; PRT; 365 AA.
AC P41221;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE WNT-5A PROTEIN PRECURSOR.
GN WNT5A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94116991; PubMed=8288227;
RA Clark C.C., Cohen I.R., Eichstetter I., Cannizzaro L.A.,
RA McPherson J.D., Wasmuth J.J., Iozzo R.V.;
RT "Molecular cloning of the human proto-oncogene Wnt-5A and mapping of
RT the gene (WNT5A) to chromosome 3p14-p21."
RT Genomics 18:249-260(1993).
CC -1- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING
CC MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF
CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.
CC -1- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE
CC EXTRACELLULAR MATRIX.
CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L20861; AAA16842.1; -
DR MIM; 164975; -
DR InterPro; IPR000970; Wnt1.
DR Pfam; PF00110; wnt; 1.
DR SMART; SM00097; WNT1; 1.
DR PROSITE; PS00246; WNT1; 1.
KW Developmental protein; Glycoprotein; Signal.
FT SIGNAL 1 22
FT CHAIN 23 365
FT CARBOHYD 99 99
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 365 AA; 40886 MW; 1B869F60D53D583B CRC64;

Query Match 50.7%; Score 36; DB 1; Length 365;
Best Local Similarity 54.5%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 RPK--PQQWF 9
   ||| | : | |
Db 154 RPKDLPDRLW 164

RESULT 78
CYB_ASTPE STANDARD; PRT; 379 AA.
AC Q33818;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME B.
GN COB OR CYTB.
OS Asterina pectinifera (Starfish).
OG Mitochondrion.
CC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Asterozoa;
CC Asteroidea; Valvatacea; Valvatida; Asterinidae; Asterina.
CC NCBI_TaxID=7594;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=95402698; PubMed=7672576;
RA Asakawa S., Himeno H., Miura K.-I., Watanabe K.;
RT "Nucleotide sequence and gene organization of the starfish Asterina
RT pectinifera mitochondrial genome.";
RL Genetics 140:1047-1060(1995).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN.
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; D16387; BAA03877.1; ALT_INIT.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c; 1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_OO; 1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 183 183 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
SQ SEQUENCE 379 AA; 42646 MW; 33A96455922F929E CRC64;

Query Match 50.7%; Score 36; DB 1; Length 379;
Best Local Similarity 54.5%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 RPKPQQWF 11
   || | |||
Db 319 RPKSQSLFW 329
```

```
RESULT 79
WN5A_MOUSE STANDARD; PRT; 379 AA.
ID P22725;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE WNT-5A PROTEIN PRECURSOR.
GN WNT5A OR WNT-5A.
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91122634; PubMed=2279700;
RA Gavin B.J., McMahon J.A., McMahon A.P.;
RT "Expression of multiple novel Wnt-1/int-1-related genes during fetal
RT and adult mouse development.";
RL Genes Dev. 4:2319-2332(1990).
CC -!- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING
CC MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF
CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.
CC -!- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE
CC EXTRACELLULAR MATRIX.
CC -!- SIMILARITY: BELONGS TO THE WNT FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M89798; AAA40567.1; -.
DR PIR; D36470; D36470.
DR MGD; MGI:98958; Wnt5a.
DR InterPro; IPR000970; Wnt1.
DR Pfam; PF00110; wnt; 1.
DR SMART; SM00097; WNT1; 1.
DR PROSITE; PS00246; WNT1; 1.
KW Developmental protein; Glycoprotein; Signal.
FT SIGNAL 1 37 POTENTIAL.
FT CHAIN 38 379 WNT-5A PROTEIN.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 379 AA; 42153 MW; E26266F47E169B50 CRC64;

Query Match 50.7%; Score 36; DB 1; Length 379;
Best Local Similarity 54.5%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 RPK--PQQWF 9
   ||| | : | |
Db 168 RPKDLPDRLW 178

RESULT 80
WN5A_RAT STANDARD; PRT; 379 AA.
ID Q9QXQ7;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
```

```
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE WNT-5A PROTEIN PRECURSOR.
GN WNT5A OR WNT-5A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Corpus luteum;
RA Lecher M.D., Walther P.R., Lareu R.R., Dharmarajan A.M., Frills R.R.;
RT "Coexpression of Wnt-5a, Wnt-4, frizzled-4, and DDC4 (frp4p, a
RT secreted frizzled, AF012891) in the pregnant ovary.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING
CC MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF
CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.
CC -1- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE
CC EXTRACELLULAR MATRIX.
CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF188333; AAF15588.1; -.
DR InterPro: IPR000970; Wnt1.
DR Pfam; PF00110; wnt; 1.
DR SMART; SM00097; WNT1; 1.
DR PROSITE; PS00246; WNT1; 1.
KW Developmental protein; Glycoprotein; Signal.
FT SIGNAL 1 37
FT CHAIN 38 379 WNT-5A PROTEIN.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 120 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 379 AA; 42267 MW; 5AAFLD22B5E3EA5E CRC64;
SQ
Query Match 50.7%; Score 36; DB 1; Length 379;
Best Local Similarity 54.5%; Pred. No. 2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;
QY 1 RPK--POOWFW 9
Db 168 RPKDLPDRLW 178
RESULT 81
CYB_BRAFL STANDARD; PRT; 380 AA.
AC 047431;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME B.
GN COB OR CYTB.
OS Branchiostoma floridae (Florida lancelet) (Amphioxus).
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7739;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-99261652; Pubmed-10331267;
RA Boore J.L., Daehler L.L., Brown W.M.;
RT "Complete sequence, gene arrangement, and genetic code of
RT mitochondrial DNA of the cephalochordate Branchiostoma floridae
(Rattus norvegicus) (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Corpus luteum;
RA Lecher M.D., Walther P.R., Lareu R.R., Dharmarajan A.M., Frills R.R.;
RT "Coexpression of Wnt-5a, Wnt-4, frizzled-4, and DDC4 (frp4p, a
RT secreted frizzled, AF012891) in the pregnant ovary.";
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING
CC MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF
CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.
CC -1- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE
CC EXTRACELLULAR MATRIX.
CC -1- SIMILARITY: BELONGS TO THE WNT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF098298; AAB88000.1; -.
DR InterPro: IPR000179; Cyt.b.b6.
DR Pfam; PF00032; cytochrome_b6_c1.
DR Pfam; PF00033; cytochrome_b6_n1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_OO; UNKNOWN_1.
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;
KW Heme.
FT METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).
FT METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).
FT METAL 183 183 IRON 2 (HEME B562 AXIAL LIGAND).
FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).
FT SEQUENCE 380 AA; 42628 MW; 3B8529CEDF83818 CRC64;
SQ
Query Match 50.7%; Score 36; DB 1; Length 380;
Best Local Similarity 63.6%; Pred. No. 2e+02;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 RPKPQOWFWLW 11
Db 319 RPLAQVLFWLM 329
RESULT 82
CYB_BRAFL STANDARD; PRT; 380 AA.
AC P92472; O79415;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME B.
GN COB OR CYTB.
OS Branchiostoma lanceolatum (Common lancelet) (Amphioxus).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
OC Branchiostoma.
OX NCBI_TaxID=7740;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE-98292550; Pubmed-9628930;
RA Spruyt N., Delarbre C., Gachelin G., Laudet V.;
RT "Complete sequence of the amphioxus (Branchiostoma lanceolatum)
RT mitochondrial genome: relations to vertebrates.";
RL Nucleic Acids Res. 26:3279-3285(1998).
RN [2]
RP SEQUENCE OF 186-380 FROM N.A.
RA Gachelin G.;
RC Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS.
```



CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
CC BOUND TO THE PROTEIN.  
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; Y16474; CAA76246.1; -;  
DR EMBL; Y09849; CAA70979.1; -;  
DR InterPro: IPR000179; Cyt\_b\_b6.  
DR Pfam; PF00032; cytochrome\_b\_c; 1.  
DR Pfam; PF00033; cytochrome\_b\_n; 1.  
DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.  
DR PROSITE; PS00193; CYTOCHROME\_B\_OO; 1.  
KW Electron transport; Mitochondrion; Respiratory chain; Transmembrane;  
KW Heme.  
FT METAL 84 84 IRON 1 (HEME B562 AXIAL LIGAND).  
FT METAL 98 98 IRON 2 (HEME B566 AXIAL LIGAND).  
FT METAL 183 183 IRON 2 (HEME B562 AXIAL LIGAND).  
FT METAL 197 197 IRON 1 (HEME B566 AXIAL LIGAND).  
FT CONFLICT 247 247 S -> C (IN REF. 2).  
SQ SEQUENCE 380 AA; 42639 MW; 514EA8E40537E814 CRC64;

Query Match 50.7%; Score 36; DB 1; Length 380;  
Best Local Similarity 63.6%; Pred. NO. 2e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 1 RPKQQQWFLM 11  
||| |  
Db 319 RPLAQVLFWM 329

RESULT 83  
WNSA\_XENLA STANDARD; PRT; 380 AA.  
ID P31286;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE WNT-5A PROTEIN PRECURSOR (XWNT-5A).  
GN WNT-5A.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Oocyte;  
RX MEDLINE=94102052; PubMed=8275867;  
RA Moon R.T., Campbell R.M., Christian J.L., McGrew L.L., Shih J.,  
RA Fraser S.;  
RT "Wnt-5A: a maternal Wnt that affects morphogenetic movements after  
RT overexpression in embryos of Xenopus laevis.";  
RL Development 119:97-111(1993).  
RN [2]  
RP SEQUENCE OF 238-363 FROM N.A.  
RC TISSUE=Embryo;  
RX MEDLINE=91122437; PubMed=1991549;  
RA Christian J.L., Gavin B.J., McMahon A.P., Moon R.T.;  
RT "Isolation of cDNAs partially encoding four Xenopus  
RT Wnt-1/int-1-related proteins and characterization of their transient  
RT expression during embryonic development.";  
RL Dev. Biol. 143:230-234(1991).  
CC -!- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALING

CC MOLECULE WHICH AFFECTS THE DEVELOPMENT OF DISCRETE REGIONS OF  
CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS.  
CC AFFECTS MORPHOGENETIC MOVEMENT.  
CC -!- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE  
CC EXTRACELLULAR MATRIX.  
CC -!- TISSUE SPECIFICITY: FOUND PRIMARILY IN ECTODERM WITH LOWER LEVELS  
CC OF EXPRESSION IN MESODERM. DETECTED IN THE HEAD AND TAIL WITH  
CC LOWER EXPRESSION IN THE MIDDLE OF THE EMBRYO. NO EXPRESSION WAS  
CC FOUND IN THE NOTOCHORD.  
CC -!- DEVELOPMENTAL STAGE: PRESENT IN OOCYTES. LEVELS DECREASE DURING  
CC EARLY EMBRYO DEVELOPMENT AND THEN INCREASE CONSIDERABLY IN  
CC NEURULA AND TADPOLE STAGES.  
CC -!- SIMILARITY: BELONGS TO THE WNT FAMILY.  
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; L19716; AAA16628.1; -;  
DR EMBL; M55056; AAA50011.1; -;  
DR InterPro: IPR000970; Wnt1.  
DR Pfam; PF00110; wnt1; 1.  
DR SMART; SM00097; Wnt1; 1.  
DR PROSITE; PS00246; WNT1; 1.  
KW Developmental protein; Glycoprotein; Signal.  
FT SIGNAL 1 40 POTENTIAL.  
FT CHAIN 41 380 WNT-5A PROTEIN.  
FT CARBOHYD 114 114 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 120 120 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 312 312 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 326 326 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CONFLICT 261 261 MISSING (IN REF. 2).  
FT CONFLICT 265 265 H -> L (IN REF. 2).  
FT CONFLICT 274 274 G -> A (IN REF. 2).  
SQ SEQUENCE 380 AA; 42519 MW; 822E2259739E815D CRC64;

Query Match 50.7%; Score 36; DB 1; Length 380;  
Best Local Similarity 54.5%; Pred. NO. 2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;  
Qy 1 RPK--PQQRWF 9  
||| |  
Db 169 RPKDLPDRLW 179

RESULT 84  
CYB\_CARPL STANDARD; PRT; 381 AA.  
ID CYB\_CARPL  
AC P34866;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CYTOCHROME B.  
GN MTCYB OR COB OR CYTB.  
OS Carcharhinus plumbeus (Sandbar shark).  
OC Mitochondrion.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;  
OC Carcharhinidae; Carcharhinus.  
OX NCBI\_TaxID=7808;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92252907; PubMed=1579163;  
RA Martin A.P., Naylor G.J.P., Palumbi S.R.;  
RT "Rates of mitochondrial DNA evolution in sharks are slow compared  
RT with mammals.";  
RL Nature 357:153-155(1992).  
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE





RL Nat. Neurosci. 2:422-426(1999).  
CC -1- FUNCTION: VOLTAGE INSENSITIVE, INSTANTANEOUS, OUTWARDLY RECTIFYING  
CC POTASSIUM CHANNEL. OUTWARD RECTIFICATION IS REVERSED AT HIGH  
CC EXTERNAL K<sup>+</sup> CONCENTRATIONS.  
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 (SHOWN HERE) AND  
CC 2/TRAAT/TRUNCATED; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, SPINAL CORD AND EYE. NOT  
CC DETECTED IN HEART, SKELETAL MUSCLE, LIVER, LUNGS, KIDNEY AND  
CC TESTIS.  
CC -1- MISCELLANEOUS: ACTIVATED BY ARACHIDONIC ACID AND OTHER UNSATURATED  
CC FATTY ACIDS. NOT AFFECTED BY VOLATILE GENERAL ANAESTHETICS SUCH AS  
CC CHLOROFORM, DIETHYL ETHER, HALOTHANE AND ISOFLURANE.  
CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM  
CC CHANNELS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF056492; AAC40181.1; -.  
CC DR MGD; MGI:1298234; Kcnk4.  
CC DR InterPro: IPR003280; 2porek\_channel.  
CC DR InterPro: IPR001622; TWIK\_channel.  
CC DR Pfam; PF02034; TWIK\_channel; 1.  
CC DR PRINTS; PR01333; 2POREKCHANEL.  
CC KW Ionic channel; Transmembrane; Ion transport; Potassium transport;  
CC Glycoprotein; Alternative splicing.  
CC FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).  
CC TRANSMEM 4 24 POTENTIAL.  
CC FT DOMAIN 89 113 PORE-FORMING (POTENTIAL).  
CC TRANSMEM 119 139 POTENTIAL.  
CC FT DOMAIN 140 171 CYTOPLASMIC (POTENTIAL).  
CC TRANSMEM 172 192 POTENTIAL.  
CC FT DOMAIN 198 222 PORE-FORMING (POTENTIAL).  
CC TRANSMEM 235 255 POTENTIAL.  
CC FT DOMAIN 256 398 CYTOPLASMIC (POTENTIAL).  
CC CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT VARSPLIC 63 67 KLIVE -> KAWAI (IN ISOFORM 2).  
CC FT VARSPLIC 68 398 MISSING (IN ISOFORM 2).  
CC SEQUENCE 398 AA; 43051 MW; 478A834B7B/AEC92 CRC64;  
  
Query Match 50.7%; Score 36; DB 1; Length 398;  
Best Local Similarity 44.4%; Pred. No. 2e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 3 KPQQWFWM 11  
:| | | | |  
DB 233 QPLVWFWM 241  
  
RESULT 89  
CIST\_SCHPO STANDARD; PRT; 473 AA.  
AC Q10306;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROBABLE CITRATE SYNTHASE, MITOCHONDRIAL PRECURSOR (EC 4.1.3.7).  
GN SPAC6C3.03.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
CC NCBI\_TaxID=4896;

RN SEQUENCE FROM N.A.  
RP STRAIN=972;  
RA Devlin K., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: CITRATE + COA = ACETYL-COA + H(2)O +  
CC OXALOACETATE.  
CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.  
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL MATRIX (BY SIMILARITY).  
CC -1- MISCELLANEOUS: CITRATE SYNTHASE IS FOUND IN NEARLY ALL CELLS  
CC CAPABLE OF OXIDATIVE METABOLISM.  
CC -1- SIMILARITY: BELONGS TO THE CITRATE SYNTHASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z69731; CA93617.2; -.  
CC DR HSSP; P23007; 5CSC.  
CC DR InterPro: IPR002020; Citrate\_synt.  
CC DR Pfam; PF00285; citrate\_synt; 1.  
CC DR PRINTS; PR00143; CITRITSNTHASE.  
CC DR PROSITE; PS00480; CITRATE\_SYNTHASE; 1.  
CC KW Hypothetical protein; Lyase; Tricarboxylic acid cycle; Mitochondrion;  
CC Transit peptide.  
CC FT TRANSIT 1 35 MITOCHONDRION (BY SIMILARITY).  
CC FT CHAIN 36 473 PROBABLE CITRATE SYNTHASE.  
CC FT ACT\_SITE 310 310 BY SIMILARITY.  
CC FT ACT\_SITE 356 356 BY SIMILARITY.  
CC FT ACT\_SITE 411 411 BY SIMILARITY.  
CC SQ SEQUENCE 473 AA; 52973 MW; 86E35E34ADC3E060 CRC64;  
  
Query Match 50.7%; Score 36; DB 1; Length 473;  
Best Local Similarity 45.5%; Pred. No. 2.4e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKPQQWFWM 11  
:| | | | |  
DB 122 QPLPESLFWLL 132  
  
RESULT 90  
CPF4\_RAT  
ID CPF4\_RAT STANDARD; PRT; 522 AA.  
AC P51869;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE CYTOCHROME P450 4F4 (EC 1.14.14.1) (CYP1F4).  
GN CYP4F4.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
CC NCBI\_TaxID=10116;  
CC  
CC SEQUENCE FROM N.A.  
RP STRAIN=SPRAGUE-DAWLEY; TISSUE=Brain;  
RX MEDLINE=96125358; PubMed=8554568;  
RA Kawashima H., Strobel H.W.;  
RT "cDNA cloning of three new forms of rat brain cytochrome P450  
RT belonging to the CYP4F subfamily";  
RL Biochem. Biophys. Res. Commun. 217:1137-1144(1995).  
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN LIVER AND KIDNEY. LOWER  
CC EXPRESSION IN BRAIN.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.

```
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U39206; AAC52358.1; -
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00464; EP45011.
DR Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum.
FT BINDING 468 468 HEME (BY SIMILARITY).
SQ SEQUENCE 522 AA; 60049 MW; 193EABE7CE8864D6 CRC64;

Query Match 50.7%; Score 36; DB 1; Length 522;
Best Local Similarity 40.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PKPQWFLM 11
Db 52 PQPKWNNFL 61

RESULT 91
DPOG_MOUSE STANDARD; PRT; 1238 AA.
AC P54099;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE DNA POLYMERASE GAMMA (EC 2.7.7.7) (MITOCHONDRIAL DNA POLYMERASE
DE CATALYTIC SUBUNIT).
GN POLG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C; TISSUE=Muscle;
RA Chang S.W., Colvin S., Sarkos P., Denniger G., Zassenhaus H.P.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN THE REPLICATION OF MITOCHONDRIAL DNA.
CC -!- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE -
CC N PYROPHOSPHATE + DNA(N).
CC -!- COFACTOR: MAGNESIUM.
CC -!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -!- MISCELLANEOUS: IN EUKARYOTES THERE ARE FIVE DNA POLYMERASES:
CC ALPHA, BETA, GAMMA, DELTA, AND EPSILON WHICH ARE RESPONSIBLE FOR
CC DIFFERENT REACTIONS OF DNA SYNTHESIS.
CC -!- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U53584; AAA98977.1; -
DR MGD; MGI:1196389; Polg.
DR InterPro; IPR002297; DNA_polg.
DR InterPro; IPR001098; DNA_pol_A.
DR InterPro; IPR001179; FKBP_PPIase.
```

```
DR Pfam; PF00476; DNA_pol_A; 1.
DR Pfam; PF00254; FKBP; 1.
DR PRINTS; PR00867; DNAPOLG.
DR SMART; SM00482; POLAC; 1.
DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
KW Transferrase; DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Mitochondrion; Magnesium.
FT DOMAIN 444 447 POLY-VAL.
FT DOMAIN 516 519 POLY-GLU.
SQ SEQUENCE 1238 AA; 138933 MW; FAF393CC5850AD9B CRC64;

Query Match 50.7%; Score 36; DB 1; Length 1238;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PKPQWFW 9
Db 147 PEPKSWAW 154

RESULT 92
Y056_HUMAN STANDARD; PRT; 1507 AA.
ID Y056_HUMAN
AC P42695;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL PROTEIN KIAA0056 (FRAGMENT).
GN KIAA0056.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96051398; PubMed=7584044;
RA Nomura N., Nagase T., Miyajima N., Suzuki T., Tanaka A., Sato S.,
RA Seki N., Kawarabayashi Y., Ishikawa K.-I., Tabata S.;
RT "Prediction of the coding sequences of unidentified human genes. II.
RT The coding sequences of 40 new genes (KIAA0041-KIAA0080) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 1:223-229(1994).
RN [2]
RP SEQUENCE OF 1171-1507 FROM N.A.
RC TISSUE=Brain;
RA Yu W.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; D29954; BAA06224.1; -
DR EMBL; AF070553; AAC28639.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 1507 AA; 169717 MW; D56BCA948FE927D2 CRC64;

Query Match 50.7%; Score 36; DB 1; Length 1507;
Best Local Similarity 83.3%; Pred. No. 6.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PKPQW 7
Db 1474 PQPQW 1479
```

Search completed: April 1, 2002, 16:20:19  
Job time: 165 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2002, 16:20:04 ; Search time 35.93 Seconds

(without alignments)  
44.781 Million cell updates/sec

Title: US-09-988-792-2

Perfect score: 71

Sequence: 1 RKPKQWFWM 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters: 244

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 50%

Maximum Match 100%

Listing first 1000 summaries

Database :

SPTREMBL\_17:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_protist:\*

12: sp\_virus:\*

13: sp Vertebrate:\*

14: sp\_unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-------|-------------|
| 1          | 48    | 67.6          | 72     | 4     | Q9Y494      |
| 2          | 48    | 67.6          | 114    | 6     | Q97947      |
| 3          | 48    | 67.6          | 128    | 4     | Q9Y6V5      |
| 4          | 48    | 67.6          | 129    | 6     | Q97948      |
| 5          | 45    | 63.4          | 453    | 2     | Q9PL30      |
| 6          | 45    | 63.4          | 455    | 2     | Q84013      |
| 7          | 44    | 62.0          | 97     | 11    | Q9Z0K2      |
| 8          | 44    | 62.0          | 115    | 11    | Q9Z0K1      |
| 9          | 44    | 62.0          | 130    | 11    | Q9Z0K0      |
| 10         | 43    | 60.6          | 314    | 2     | Q9A645      |
| 11         | 43    | 60.6          | 365    | 2     | Q9JY60      |
| 12         | 43    | 60.6          | 365    | 2     | Q9JY44      |
| 13         | 43    | 60.6          | 818    | 12    | Q9PWS3      |
| 14         | 43    | 60.6          | 898    | 12    | Q9PYP2      |
| 15         | 41    | 57.7          | 158    | 2     | Q34889      |
| 16         | 41    | 57.7          | 159    | 2     | Q84560      |
| 17         | 41    | 57.7          | 209    | 11    | Q9DAS6      |
| 18         | 41    | 57.7          | 261    | 2     | Q9HXA6      |
| 19         | 41    | 57.7          | 286    | 2     | Q9PAV0      |

|    |      |      |      |    |        |                    |
|----|------|------|------|----|--------|--------------------|
| 20 | 41   | 57.7 | 290  | 2  | Q9AKK5 | Q9akk5 rickettsia  |
| 21 | 41   | 57.7 | 291  | 2  | Q55507 | Q55507 synechocyst |
| 22 | 41   | 57.7 | 304  | 1  | Q9HNN1 | Q9hnn1 halobacteri |
| 23 | 41   | 57.7 | 392  | 2  | Q9F849 | Q9f849 streptomyce |
| 24 | 41   | 57.7 | 529  | 2  | P74332 | P74332 synechocyst |
| 25 | 40   | 56.3 | 45   | 5  | Q9V6W6 | Q9v6w6 drosophila  |
| 26 | 40   | 56.3 | 154  | 11 | Q9R129 | Q9r129 rattus norv |
| 27 | 40   | 56.3 | 155  | 11 | Q9R132 | Q9r132 rattus norv |
| 28 | 40   | 56.3 | 155  | 11 | Q9R131 | Q9r131 rattus norv |
| 29 | 40   | 56.3 | 155  | 11 | Q9R127 | Q9r127 rattus norv |
| 30 | 40   | 56.3 | 155  | 11 | Q9JKI6 | Q9jki6 mus saxicol |
| 31 | 40   | 56.3 | 273  | 2  | Q9KRO9 | Q9kr9 vibrio chol  |
| 32 | 40   | 56.3 | 289  | 2  | Q9JYV2 | Q9jyv2 neisseria m |
| 33 | 40   | 56.3 | 289  | 2  | Q9JYU4 | Q9jyu4 neisseria m |
| 34 | 40   | 56.3 | 290  | 2  | Q9ZCL1 | Q9zcl1 rickettsia  |
| 35 | 40   | 56.3 | 290  | 2  | Q9AKF0 | Q9akf0 rickettsia  |
| 36 | 40   | 56.3 | 290  | 2  | Q9AKA6 | Q9aka6 rickettsia  |
| 37 | 40   | 56.3 | 318  | 2  | Q9KVD3 | Q9kvd3 vibrio chol |
| 38 | 40   | 56.3 | 332  | 2  | Q82937 | Q82937 escherichia |
| 39 | 40   | 56.3 | 343  | 2  | Q9ZGU3 | Q9zgu3 escherichia |
| 40 | 40   | 56.3 | 363  | 2  | P73727 | P73727 synechocyst |
| 41 | 40   | 56.3 | 367  | 10 | Q9C5G0 | Q9c5g0 arabidopsis |
| 42 | 40   | 56.3 | 381  | 10 | Q9FYI5 | Q9fyi5 arabidopsis |
| 43 | 40   | 56.3 | 411  | 4  | Q9NRT2 | Q9nrt2 homo sapien |
| 44 | 40   | 56.3 | 639  | 4  | Q9H8K3 | Q9h8k3 homo sapien |
| 45 | 40   | 56.3 | 665  | 5  | Q9V7L5 | Q9v7l5 drosophila  |
| 46 | 40   | 56.3 | 774  | 4  | Q9NWY0 | Q9nwy0 homo sapien |
| 47 | 40   | 56.3 | 797  | 10 | Q9CAX2 | Q9cax2 arabidopsis |
| 48 | 40   | 56.3 | 937  | 3  | Q9HE57 | Q9he57 neosporea   |
| 49 | 39.5 | 55.6 | 548  | 2  | Q9RV51 | Q9rv51 deinococcus |
| 50 | 39   | 54.9 | 55   | 8  | Q79563 | Q79563 mustelus ma |
| 51 | 39   | 54.9 | 99   | 12 | Q69572 | Q69572 human herpe |
| 52 | 39   | 54.9 | 208  | 2  | Q55572 | Q55572 synechocyst |
| 53 | 39   | 54.9 | 341  | 2  | Q9JSJ3 | Q9jsj3 chlamydia p |
| 54 | 39   | 54.9 | 361  | 4  | Q9PLW6 | Q9plw6 homo sapien |
| 55 | 39   | 54.9 | 383  | 2  | Q9F669 | Q9f669 pseudomonas |
| 56 | 39   | 54.9 | 412  | 2  | Q9ZJ55 | Q9zj55 helicobacte |
| 57 | 39   | 54.9 | 416  | 2  | Q9HU99 | Q9hu99 pseudomonas |
| 58 | 39   | 54.9 | 452  | 2  | Q9PIU6 | Q9piu6 campylobact |
| 59 | 39   | 54.9 | 462  | 2  | Q9K217 | Q9k217 chlamydia p |
| 60 | 39   | 54.9 | 467  | 2  | Q9Z983 | Q9z983 chlamydia p |
| 61 | 39   | 54.9 | 480  | 1  | Q9YDN6 | Q9ydn6 aeropyrum p |
| 62 | 39   | 54.9 | 499  | 5  | Q9VJ81 | Q9vj81 drosophila  |
| 63 | 39   | 54.9 | 524  | 4  | Q9HB16 | Q9hbi6 homo sapien |
| 64 | 39   | 54.9 | 525  | 2  | Q83145 | Q83145 treponema p |
| 65 | 39   | 54.9 | 778  | 13 | Q9PDA3 | Q9pda3 xenopus lae |
| 66 | 39   | 54.9 | 825  | 13 | Q9I9Q6 | Q9i9q6 xenopus lae |
| 67 | 39   | 54.9 | 827  | 5  | Q9U6M1 | Q9u6m1 leishmania  |
| 68 | 39   | 54.9 | 852  | 13 | Q9IAI1 | Q9iai1 gallus gall |
| 69 | 39   | 54.9 | 853  | 13 | Q9W6Q2 | Q9w6q2 gallus gall |
| 70 | 39   | 54.9 | 854  | 13 | Q93244 | Q93244 oncorhynch  |
| 71 | 39   | 54.9 | 854  | 13 | Q91953 | Q91953 coturnix co |
| 72 | 39   | 54.9 | 862  | 11 | Q9WVS9 | Q9wvs9 rattus norv |
| 73 | 39   | 54.9 | 875  | 13 | Q9W7C3 | Q9w7c3 gallus gall |
| 74 | 39   | 54.9 | 893  | 13 | Q9W6J4 | Q9w6j4 brachydanio |
| 75 | 39   | 54.9 | 1023 | 5  | Q9VSB0 | Q9vsb0 drosophila  |
| 76 | 39   | 54.9 | 1219 | 2  | Q53785 | Q53785 streptomyce |
| 77 | 38.5 | 54.2 | 673  | 10 | Q9FVG4 | Q9fvq4 zea mays (m |
| 78 | 38   | 53.5 | 105  | 2  | Q9HZ24 | Q9hz24 pseudomonas |
| 79 | 38   | 53.5 | 122  | 11 | Q62024 | Q62024 mus musculu |
| 80 | 38   | 53.5 | 124  | 11 | Q62023 | Q62023 mus musculu |
| 81 | 38   | 53.5 | 175  | 8  | Q9T6A9 | Q9t6a9 culicoides  |
| 82 | 38   | 53.5 | 188  | 11 | Q9D1Z1 | Q9d1z1 mus musculu |
| 83 | 38   | 53.5 | 220  | 2  | Q07878 | Q07878 sphingomona |
| 84 | 38   | 53.5 | 253  | 2  | Q9ANC0 | Q9anc0 bradyrhizob |
| 85 | 38   | 53.5 | 260  | 5  | Q9GUP1 | Q9gup1 caenorhabdi |
| 86 | 38   | 53.5 | 270  | 6  | P79391 | P79391 bos taurus  |
| 87 | 38   | 53.5 | 273  | 4  | P78380 | P78380 homo sapien |
| 88 | 38   | 53.5 | 274  | 6  | Q9TTK7 | Q9ttk7 sus scrofa  |
| 89 | 38   | 53.5 | 322  | 11 | Q99L65 | Q99l65 mus musculu |
| 90 | 38   | 53.5 | 330  | 11 | Q62022 | Q62022 mus musculu |
| 91 | 38   | 53.5 | 356  | 13 | Q98952 | Q98952 gallus gall |
| 92 | 38   | 53.5 | 363  | 11 | Q9EQ09 | Q9eq09 mus musculu |

|     |      |      |      |    |        |                     |     |      |      |      |    |         |
|-----|------|------|------|----|--------|---------------------|-----|------|------|------|----|---------|
| 93  | 38   | 53.5 | 364  | 11 | 070156 | 070156 rattus norv  | 166 | 37   | 52.1 | 1589 | 13 | Q91588  |
| 94  | 38   | 53.5 | 368  | 2  | Q9A8W6 | Q9A8W6 xyllobacter  | 167 | 37   | 52.1 | 1635 | 5  | O17368  |
| 95  | 38   | 53.5 | 373  | 2  | Q9PDR6 | Q9PDR6 xylella fas  | 168 | 37   | 52.1 | 2458 | 2  | O51827  |
| 96  | 38   | 53.5 | 392  | 10 | Q9FHER | Q9FHER arabidopsis  | 169 | 36.5 | 51.4 | 523  | 11 | Q99N19  |
| 97  | 38   | 53.5 | 401  | 11 | Q9JIK2 | Q9JIK2 cricetus     | 170 | 36.5 | 51.4 | 523  | 11 | Q99KY6  |
| 98  | 38   | 53.5 | 412  | 2  | O26064 | O26064 helicobacte  | 171 | 36   | 50.7 | 93   | 10 | Q9SDN8  |
| 99  | 38   | 53.5 | 421  | 10 | Q9CAP2 | Q9CAP2 arabidopsis  | 172 | 36   | 50.7 | 103  | 2  | P73564  |
| 100 | 38   | 53.5 | 433  | 11 | Q99N88 | Q99N88 rattus norv  | 173 | 36   | 50.7 | 104  | 5  | O76548  |
| 101 | 38   | 53.5 | 455  | 10 | O65524 | O65524 arabidopsis  | 174 | 36   | 50.7 | 105  | 2  | Q9AF27  |
| 102 | 38   | 53.5 | 458  | 5  | Q9WLD1 | Q9WLD1 drosophila   | 175 | 36   | 50.7 | 125  | 11 | Q9DCS0  |
| 103 | 38   | 53.5 | 635  | 5  | Q9VUK7 | Q9VUK7 drosophila   | 176 | 36   | 50.7 | 135  | 8  | Q36742  |
| 104 | 38   | 53.5 | 684  | 5  | Q9UAC1 | Q9UAC1 leishmania   | 177 | 36   | 50.7 | 154  | 2  | Q36RVC1 |
| 105 | 38   | 53.5 | 691  | 5  | Q9UAB7 | Q9UAB7 leishmania   | 178 | 36   | 50.7 | 171  | 2  | Q9HWX8  |
| 106 | 38   | 53.5 | 697  | 10 | Q9LUQ4 | Q9LUQ4 arabidopsis  | 179 | 36   | 50.7 | 241  | 10 | Q9FRE0  |
| 107 | 38   | 53.5 | 698  | 5  | Q9UAC0 | Q9UAC0 leishmania   | 180 | 36   | 50.7 | 250  | 8  | Q9TAK3  |
| 108 | 38   | 53.5 | 700  | 5  | Q9UAB9 | Q9UAB9 leishmania   | 181 | 36   | 50.7 | 254  | 6  | O19110  |
| 109 | 38   | 53.5 | 722  | 10 | Q9WXL1 | Q9WXL1 arabidopsis  | 182 | 36   | 50.7 | 268  | 5  | O21187  |
| 110 | 38   | 53.5 | 901  | 12 | Q9DXA2 | Q9DXA2 choristoneu  | 183 | 36   | 50.7 | 280  | 2  | Q9KRT5  |
| 111 | 38   | 53.5 | 909  | 2  | P74693 | P74693 synecocyst   | 184 | 36   | 50.7 | 287  | 2  | Q9KX15  |
| 112 | 38   | 53.5 | 956  | 5  | O00908 | O00908 cryptospori  | 185 | 36   | 50.7 | 302  | 2  | O53945  |
| 113 | 38   | 53.5 | 1184 | 4  | O75339 | O75339 homo sapien  | 186 | 36   | 50.7 | 320  | 11 | Q9CYX6  |
| 114 | 38   | 53.5 | 1216 | 11 | Q9QYV8 | Q9QYV8 rattus norv  | 187 | 36   | 50.7 | 343  | 8  | Q9B4H0  |
| 115 | 38   | 53.5 | 1216 | 11 | Q9QYV7 | Q9QYV7 rattus norv  | 188 | 36   | 50.7 | 347  | 2  | O47471  |
| 116 | 38   | 53.5 | 1528 | 5  | P81137 | P81137 manduca sex  | 189 | 36   | 50.7 | 359  | 8  | Q9MM76  |
| 117 | 38   | 53.5 | 1717 | 5  | Q9GPJ9 | Q9GPJ9 manduca sex  | 190 | 36   | 50.7 | 359  | 4  | Q9BV04  |
| 118 | 38   | 53.5 | 1832 | 5  | O96503 | O96503 cryptospori  | 191 | 36   | 50.7 | 360  | 6  | Q9XT34  |
| 119 | 38   | 53.5 | 2237 | 5  | Q9V122 | Q9V122 drosophila   | 192 | 36   | 50.7 | 362  | 11 | O35886  |
| 120 | 37.5 | 52.8 | 100  | 14 | Q99IT0 | Q99IT0 uncultured   | 193 | 36   | 50.7 | 362  | 11 | Q9R220  |
| 121 | 37.5 | 52.8 | 213  | 2  | Q99ZD5 | Q99ZD5 streptococ   | 194 | 36   | 50.7 | 362  | 11 | Q9R219  |
| 122 | 37.5 | 52.8 | 300  | 11 | Q9CY68 | Q9CY68 mus musculu  | 195 | 36   | 50.7 | 369  | 2  | P73843  |
| 123 | 37.5 | 52.8 | 473  | 11 | Q9Z1X2 | Q9Z1X2 mus musculu  | 196 | 36   | 50.7 | 380  | 8  | Q9XPE2  |
| 124 | 37.5 | 52.8 | 474  | 11 | O08888 | O08888 cricetus     | 197 | 36   | 50.7 | 385  | 13 | Q9RGX6  |
| 125 | 37.5 | 52.8 | 487  | 4  | Q9BVG9 | Q9BVG9 homo sapien  | 198 | 36   | 50.7 | 418  | 11 | Q99JJ4  |
| 126 | 37.5 | 52.8 | 1177 | 12 | Q92611 | Q92611 pseudorabie  | 199 | 36   | 50.7 | 454  | 5  | Q9VNP0  |
| 127 | 37.5 | 52.8 | 1194 | 12 | Q9E1V7 | Q9E1V7 cercopithe   | 200 | 36   | 50.7 | 482  | 8  | Q9MGA9  |
| 128 | 37.5 | 52.8 | 1197 | 12 | Q99101 | Q99101 herpes simp  | 201 | 36   | 50.7 | 534  | 11 | Q99N18  |
| 129 | 37.5 | 52.8 | 1203 | 12 | Q99549 | Q99549 bovine herp  | 202 | 36   | 50.7 | 603  | 2  | Q9L217  |
| 130 | 37.5 | 52.8 | 1208 | 12 | O39273 | O39273 equine herp  | 203 | 36   | 50.7 | 605  | 2  | O86684  |
| 131 | 37   | 52.1 | 53   | 8  | Q9TBB3 | Q9TBB3 tadorna var  | 204 | 36   | 50.7 | 611  | 11 | O60850  |
| 132 | 37   | 52.1 | 53   | 8  | Q9TBB2 | Q9TBB2 tadorna tad  | 205 | 36   | 50.7 | 633  | 5  | O9VT54  |
| 133 | 37   | 52.1 | 116  | 2  | Q9RM36 | Q9RM36 escherichia  | 206 | 36   | 50.7 | 643  | 1  | O29580  |
| 134 | 37   | 52.1 | 181  | 2  | Q9ADB8 | Q9ADB8 streptomyc   | 207 | 36   | 50.7 | 656  | 10 | O65001  |
| 135 | 37   | 52.1 | 220  | 3  | O14264 | O14264 schizosacch  | 208 | 36   | 50.7 | 670  | 2  | Q9RXQ7  |
| 136 | 37   | 52.1 | 226  | 5  | Q9V424 | Q9V424 drosophila   | 209 | 36   | 50.7 | 704  | 2  | Q9KKL9  |
| 137 | 37   | 52.1 | 273  | 11 | Q9D3C8 | Q9D3C8 mus musculu  | 210 | 36   | 50.7 | 729  | 4  | Q9UJ93  |
| 138 | 37   | 52.1 | 292  | 2  | Q9AP18 | Q9AP18 methylobact  | 211 | 36   | 50.7 | 774  | 2  | Q9K2S2  |
| 139 | 37   | 52.1 | 314  | 5  | Q9VHB7 | Q9VHB7 drosophila   | 212 | 36   | 50.7 | 804  | 2  | P70811  |
| 140 | 37   | 52.1 | 331  | 3  | O00893 | Q00893 collettotric | 213 | 36   | 50.7 | 805  | 2  | Q9RGZ5  |
| 141 | 37   | 52.1 | 359  | 4  | Q9Y231 | Q9Y231 homo sapien  | 214 | 36   | 50.7 | 817  | 4  | Q9H8I4  |
| 142 | 37   | 52.1 | 359  | 11 | O88819 | O88819 mus musculu  | 215 | 36   | 50.7 | 823  | 4  | O15033  |
| 143 | 37   | 52.1 | 359  | 11 | Q9JIG1 | Q9JIG1 cricetus     | 216 | 36   | 50.7 | 829  | 2  | O55414  |
| 144 | 37   | 52.1 | 359  | 11 | Q95JB3 | Q95JB3 rattus norv  | 217 | 36   | 50.7 | 861  | 2  | O06944  |
| 145 | 37   | 52.1 | 365  | 6  | Q9TQQ3 | Q9TQQ3 bos taurus   | 218 | 36   | 50.7 | 886  | 5  | Q9VZV1  |
| 146 | 37   | 52.1 | 406  | 2  | Q9KQW4 | Q9KQW4 vibrio chol  | 219 | 36   | 50.7 | 894  | 4  | Q9C096  |
| 147 | 37   | 52.1 | 421  | 2  | Q9EX44 | Q9EX44 streptomyc   | 220 | 36   | 50.7 | 959  | 4  | Q9C096  |
| 148 | 37   | 52.1 | 467  | 2  | Q9CK08 | Q9CK08 pasteurella  | 221 | 36   | 50.7 | 1038 | 5  | O01261  |
| 149 | 37   | 52.1 | 529  | 2  | Q9FDI3 | Q9FDI3 brevibacter  | 222 | 36   | 50.7 | 1039 | 5  | O23567  |
| 150 | 37   | 52.1 | 536  | 2  | Q9KLT3 | Q9KLT3 vibrio chol  | 223 | 36   | 50.7 | 1043 | 5  | Q9VL84  |
| 151 | 37   | 52.1 | 555  | 2  | Q9RPG1 | Q9RPG1 myxococcus   | 224 | 36   | 50.7 | 1148 | 4  | Q9H6W7  |
| 152 | 37   | 52.1 | 595  | 2  | Q9RJ38 | Q9RJ38 streptomyc   | 225 | 36   | 50.7 | 1236 | 4  | Q9C012  |
| 153 | 37   | 52.1 | 604  | 8  | Q9XMR9 | Q9XMR9 tetrahymena  | 226 | 36   | 50.7 | 1239 | 11 | Q9J128  |
| 154 | 37   | 52.1 | 645  | 11 | Q62094 | Q62094 mus musculu  | 227 | 36   | 50.7 | 1243 | 5  | O9NGT8  |
| 155 | 37   | 52.1 | 656  | 12 | Q9QU36 | Q9QU36 ttv-like mi  | 228 | 36   | 50.7 | 1377 | 5  | P91854  |
| 156 | 37   | 52.1 | 669  | 13 | Q9PVX6 | Q9PVX6 cynops pyrr  | 229 | 36   | 50.7 | 1493 | 5  | Q9GPA0  |
| 157 | 37   | 52.1 | 765  | 10 | Q9SRV5 | Q9SRV5 arabidopsis  | 230 | 36   | 50.7 | 1537 | 11 | Q9JLA3  |
| 158 | 37   | 52.1 | 765  | 10 | Q9LM03 | Q9LM03 solanum tub  | 231 | 36   | 50.7 | 1555 | 4  | Q9NYU2  |
| 159 | 37   | 52.1 | 806  | 10 | Q9EFF2 | Q9EFF2 arabidopsis  | 232 | 36   | 50.7 | 1620 | 4  | Q9NUT8  |
| 160 | 37   | 52.1 | 918  | 3  | Q9HEW7 | Q9HEW7 cladosporiu  | 233 | 36   | 50.7 | 1674 | 10 | Q9FV08  |
| 161 | 37   | 52.1 | 944  | 3  | O60043 | O60043 metarhizium  | 234 | 36   | 50.7 | 1952 | 4  | O9UJ92  |
| 162 | 37   | 52.1 | 1002 | 5  | Q9W570 | Q9W570 drosophila   | 235 | 36   | 50.7 | 2135 | 4  | O43157  |
| 163 | 37   | 52.1 | 1081 | 4  | Q9H8F3 | Q9H8F3 homo sapien  | 236 | 36   | 50.7 | 2135 | 4  | Q9UIV7  |
| 164 | 37   | 52.1 | 1301 | 4  | O94987 | O94987 homo sapien  | 237 | 36   | 50.7 | 2396 | 5  | O77291  |
| 165 | 37   | 52.1 | 1418 | 13 | O93457 | O93457 scophthalmu  | 238 | 36   | 50.7 | 5201 | 5  | Q9U479  |



239 36 50.7 5385 5 Q9V6V3 Q9v6v3 drosophila  
240 36 50.7 5496 5 Q9V6V2 Q9v6v2 drosophila  
241 36 50.7 8805 5 Q9V6V4 Q9v6v4 drosophila  
242 35.5 50.0 383 2 Q56675 Q56675 vibrio chol  
243 35.5 50.0 842 13 Q9W6K7 Q9w6k7 brachydanio  
244 35.5 50.0 4861 4 Q15751 Q15751 homo sapien

## ALIGNMENTS

RESULT 1  
Q9Y494 ID Q9Y494 PRELIMINARY; PRT; 72 AA.  
AC Q9Y494;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GAMMA PREPROTACHYKININ (FRAGMENT).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BLOOD, AND BRAIN;  
RA Lai J.P., Douglas S.D., Rappaport E., Wu J.M., Ho W.Z.;  
RT "Identification of a Delta Isoform of preprotachykinin mRNA in Human  
mononuclear phagocytes and lymphocytes";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF050657; AAC15703.1; -;  
DR InterPro; IPR002040; Tachykinin.  
DR InterPro; IPR003580; Protachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR PROSITE; PS00267; TACHYKININ; UNKNOWN\_2.  
DR SMART; SM00203; TK; 2.  
FT NON\_TER 1  
FT NON\_TER 72  
SQ SEQUENCE 72 AA; 8274 MW; 2C02B2BA41EAD16 CRC64;

Query Match 67.6%; Score 48; DB 4; Length 72;  
Best Local Similarity 81.8%; Pred. No. 1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
Db 23 RPKPQQFFGLM 33  
|||||:|

RESULT 2  
Q97947 ID Q97947 PRELIMINARY; PRT; 114 AA.  
AC Q97947;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GAMMA PREPROTACHYKININ I.  
OS Tupia belangeri (northern tree shrew).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupia.  
OX NCBI\_TaxID=37347;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Heitland A., Maegert H.J., Kruhoffer M., Forssmann W.G.;  
RT "Tachykinin precursors are highly conserved among different mammals.";  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z50785; CAA90648.1; -;  
DR InterPro; IPR002040; Tachykinin.  
DR InterPro; IPR003580; Protachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR ProDom; PD005598; Protachykinin; 1.

DR PROSITE; PS00267; TACHYKININ; UNKNOWN\_2.  
DR SMART; SM00203; TK; 2.  
FT CHAIN 58 68 SUBSTANCE P.  
FT CHAIN 72 92 NEUROPEPTIDE GAMMA.  
FT CHAIN 83 92 NEUROKININ A.  
SQ SEQUENCE 114 AA; 13281 MW; B439C3D27FDA7CAB CRC64;

Query Match 67.6%; Score 48; DB 6; Length 114;  
Best Local Similarity 81.8%; Pred. No. 1.6;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
Db 58 RPKPQQFFGLM 68  
|||||:|

RESULT 3  
Q9Y6V5 ID Q9Y6V5 PRELIMINARY; PRT; 128 AA.  
AC Q9Y6V5;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE WUGSC:H-DJ0841B21.1 PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kalicki J., Angell S.;  
RT "The sequence of Homo sapiens PAC clone DJ0841B21";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Waterston R.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC004140; AAC02754.1; -;  
DR InterPro; IPR002040; Tachykinin.  
DR InterPro; IPR003580; Protachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR PROSITE; PS00267; TACHYKININ; UNKNOWN\_2.  
DR SMART; SM00203; TK; 1.  
SQ SEQUENCE 128 AA; 14770 MW; 0F8D61774AFEC1CA CRC64;

Query Match 67.6%; Score 48; DB 4; Length 128;  
Best Local Similarity 81.8%; Pred. No. 1.8;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
Db 58 RPKPQQFFGLM 68  
|||||:|

RESULT 4  
Q97948 ID Q97948 PRELIMINARY; PRT; 129 AA.  
AC Q97948;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE BETA PREPROTACHYKININ I.  
OS Tupia belangeri (northern tree shrew).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Scandentia; Tupaiidae; Tupia.  
OX NCBI\_TaxID=37347;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Heitland A., Maegert H.J., Kruhoffer M., Forssmann W.G.;  
RT "Tachykinin precursors are highly conserved among different mammals.";

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z50786; CAA90649.1; -  
DR InterPro: IPR002040; Tachykinin.  
DR InterPro: IPR003580; Protachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR ProDom; PD005598; Protachykinin; 1.  
DR PROSITE; PS00267; TACHYKININ; UNKNOWN\_2.  
DR SMART; SM00203; TK; 2.  
FT CHAIN 58 68 SUBSTANCE P.  
FT CHAIN 72 107 NEUROPEPTIDE K.  
FT CHAIN 98 107 NEUROKININ A.  
SQ SEQUENCE 129 AA; 14941 MW; 5855E7ADC2D8674E CRC64;

Query Match 67.6%; Score 48; DB 6; Length 129;

Best Local Similarity 81.8%; Pred. No. 1.8; Mismatches 1; Indels 0; Gaps 0;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11

|||||:|

Db 58 RPKPQQFFGLM 68

RESULT 5

Q9PL30

ID Q9PL30 PRELIMINARY; PRT; 453 AA.

AC Q9PL30;

DT 01-OCT-2000 (TrEMBLrel. 15, Created)

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE, PUTATIVE.

GN TC0278.

OS Chlamydia muridarum.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI\_TaxID=83560;

RP [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MOPN / NIGG;

RX MEDLINE=20150255; PubMed=10684935;

RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,

RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,

RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,

RA Gwin M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,

RA Eisen J., Fraser C.M.;

RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia

RT pneumoniae AR39.";

RL Nucleic Acids Res. 28:1397-1406(2000).

DR EMBL; AE002295; AAF39146.1; -

DR TIGR; TC0278; -

KW Complete proteome.

SQ SEQUENCE 453 AA; 51463 MW; 6221515A00A093FF CRC64;

Query Match

Best Local Similarity 63.4%; Score 45; DB 2; Length 453;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFWL 10

|||||

Db 303 KPEQWLWL 310

RESULT 6

O84013

ID O84013 PRELIMINARY; PRT; 455 AA.

AC O84013;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE ACYLTRANSFERASE.

GN HTRB OR CT010.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI\_TaxID=813;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=D/UW-3/CX;

RX MEDLINE=99000809; PubMed=9784136;

RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe F., Aravind L.,

RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,

RA Davis R.W.;

RT "Genome sequence of an obligate intracellular pathogen of humans:

RT Chlamydia trachomatis";

RL Science 282:754-759(1998).

DR EMBL; AE001275; AAC67600.1; -

KW Transferase; Acyltransferase; Complete proteome.

SQ SEQUENCE 455 AA; 52058 MW; 0404B6326C67ACCF CRC64;

Query Match

Best Local Similarity 63.4%; Score 45; DB 2; Length 455;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFWL 10

|||||

Db 303 KPEQWLWL 310

RESULT 7

Q9Z0K2

ID Q9Z0K2 PRELIMINARY; PRT; 97 AA.

AC Q9Z0K2;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE DELTA PREPROTACHYKININ I.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystricognathi; Cavidae; Cavia.

OX NCBI\_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PIRBRIDGE WHITE; TISSUE=BRAIN;

RA Heitland A., Maegert H.J., Kruehoffer M., Forssmann W.G.;

RT "Tachykinin precursors are highly conserved among different mammals.";

RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.

DR EMBL; Z50782; CAA90645.1; -

DR InterPro; IPR003580; Protachykinin.

DR ProDom; PD005598; Protachykinin; 1.

FT CHAIN 58 68 SUBSTANCE P.

SQ SEQUENCE 97 AA; 11222 MW; FFD50C3297E3F7E3 CRC64;

Query Match

Best Local Similarity 62.0%; Score 44; DB 11; Length 97;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11

|||||

Db 58 RPKPQQSFGLM 68

RESULT 8

Q9Z0K1

ID Q9Z0K1 PRELIMINARY; PRT; 115 AA.

AC Q9Z0K1;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE GAMMA PREPROTACHYKININ I.

OS Cavia porcellus (Guinea pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Hystricognathi; Cavidae; Cavia.

OX NCBI\_TaxID=10141;

RN [1]

RP SEQUENCE FROM N.A.

```

RC STRAIN=PIRBRIDGE WHITE; TISSUE=BRAIN;
RA Heitland A., Maegert H.J., Kruhooffer M., Forssmann W.G.;
RT "Tachykinin precursors are highly conserved among different mammals.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z50783; CAA90646.1; -.
DR InterPro; IPR002040; Tachykinin.
DR ProDom; PD003580; Protachykinin.
DR PROSITE; PS005598; Protachykinin; 1.
DR CHAIN 58 68 SUBSTANCE P.
FT CHAIN 72 92 NEUROPEPTIDE GAMMA.
FT CHAIN 83 92 NEUROKININ A.
SQ SEQUENCE 115 AA; 13190 MW; 39EFFB8CBB47174 CRC64;

Query Match 62.0%; Score 44; DB 11; Length 115;
Best Local Similarity 81.8%; Pred. No. 6.7;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11
||||| | | |
Db 58 RPKPQQSFGLM 68

RESULT 9
Q9Z0K0 PRELIMINARY; PRT; 130 AA.
AC Q9Z0K0;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE BETA PREPROTACHYKININ I.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
ON NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PIRBRIDGE WHITE; TISSUE=BRAIN;
RA Heitland A., Maegert H.J., Kruhooffer M., Forssmann W.G.;
RT "Tachykinin precursors are highly conserved among different mammals.";
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z50784; CAA90647.1; -.
DR InterPro; IPR002040; Tachykinin.
DR ProDom; PD003580; Protachykinin.
DR PROSITE; PS00267; TACHYKININ; UNKNOWN_1.
FT CHAIN 58 68 SUBSTANCE P.
FT CHAIN 72 107 NEUROPEPTIDE K.
FT CHAIN 98 107 NEUROKININ A.
SQ SEQUENCE 130 AA; 14850 MW; C4B2F55B6A60A7C0 CRC64;

Query Match 62.0%; Score 44; DB 11; Length 130;
Best Local Similarity 81.8%; Pred. No. 7.6;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11
||||| | | |
Db 58 RPKPQQSFGLM 68

RESULT 10
Q9A645 PRELIMINARY; PRT; 314 AA.
AC Q9A645;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN CC2250.
GN CC2250.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;

```

```

OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005895; AAK24221.1; -.
DR TIGR; CC2250; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 314 AA; 35316 MW; 59825ADEF764362A CRC64;

Query Match 60.6%; Score 43; DB 2; Length 314;
Best Local Similarity 50.0%; Pred. No. 26;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 10
| : | : | : |
Db 282 RERPAEWFV 291

RESULT 11
Q9JY60 PRELIMINARY; PRT; 365 AA.
ID Q9JY60;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CYTOCHROME C OXIDASE, SUBUNIT III.
GN NM1723.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RP SEQUENCE FROM N.A.
RX STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tetelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citterone H., Clark E.B.,
RA Cotton M.D., Uterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.";
RL Science 287:1809-1815(2000).
DR EMBL; AE002522; AAF42068.1; -.
DR TIGR; NM1723; -.
DR InterPro; IPR000345; CytC_heme_bind.
DR InterPro; IPR003088; Cyt_C1.
DR InterPro; IPR002329; Cyt_C1C.
DR Pfam; PF00034; cytochrome.c.2.
DR PRINTS; PR00605; CYTOCHROME_C.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 365 AA; 40039 MW; AF344435A51EB4A2 CRC64;

Query Match 60.6%; Score 43; DB 2; Length 365;
Best Local Similarity 56.7%; Pred. No. 30;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RPKPQQWFWM 10

```

Db 62 PLPRWFWL 70  
| | | | |

## RESULT 12

Q9JUT44 PRELIMINARY; PRT; 365 AA.  
AC Q9JUT44;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE PUTATIVE CYTOCHROME C.  
GN NMA1977.  
OS Neisseria meningitidis (serogroup A).  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=65699;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;  
RX MEDLINE=20222556; PubMed=10761919;  
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C., Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T., Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S., Jagels K., Leather S., Moule S., Mungall K., Quail M.A., Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J., Whitehead S., Spratt B.G., Barrell B.G.;  
RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491";  
RL Nature 404:502-506(2000).  
RE EMBL; AL162757; CAB85197.1; -.  
DR InterPro; IPR000345; CytC\_heme\_bind.  
DR InterPro; IPR003088; Cyt.C1.  
DR InterPro; IPR002329; Cyt.C1C.  
DR Pfam; PF00034; cytochrome\_c; 2.  
DR PRINTS; P00605; CYTOCHROME\_C1C.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 365 AA; 40011 MW; AF223552A51EB4A2 CRC64;

Query Match 60.6%; Score 43; DB 2; Length 365;

Best Local Similarity 66.7%; Pred. No. 30; Mismatches 1; Indels 2; Gaps 0;

Matches 6; Conservative 1; Indels 2; Gaps 0;  
QY 2 PKPQWFWL 10  
| | | | |

Db 62 PLPRWFWL 70

## RESULT 13

Q9PWS3 PRELIMINARY; PRT; 818 AA.  
AC Q9PWS3;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GAG-ABL PROTEIN (FRAGMENT).  
OS Abelson murine leukemia virus.  
OC Viruses; Retroviridae; Retroviridae; Mammalian type C retroviruses.  
OX NCBI\_TaxID=11789;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lee R., Paskind M., Wang J.Y.J., Baltimore D.;  
RT "Abelson (P160) murine leukemia virus (Ab-MLV) abl gene."; (In) Weiss R., Teich N., Varmus H., Coffin J. (eds.);  
RL RNA tumor viruses, pp.861-868, Cold Spring Harbor Laboratory Press, New York (1985).  
RE EMBL; X02963; CAB56204.1; -.  
DR HSP; P00519; 2ABL.  
DR InterPro; IPR00719; Euk\_pkinase.  
DR InterPro; IPR000980; SH2.  
DR InterPro; IPR001245; Tyr\_kin.  
DR Pfam; PF00069; pkinase; 1.

Pfam; PF00017; SH2; 1.  
DR PRINTS; PR00401; SH2DOMAIN.  
DR PRINTS; PR00109; TYRKINASE.  
DR SMART; SM00252; SH2; 1.  
DR SMART; SM00219; TyrKc; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.  
DR PROSITE; PS00001; SH2; 1.  
KW ATP-binding; Transferase; Tyrosine-protein kinase.  
FT NON\_TER 1  
SQ SEQUENCE 818 AA; 90973 MW; C2F5F417D0A9FE0C CRC64;

Query Match 60.6%; Score 43; DB 12; Length 818;  
Best Local Similarity 77.8%; Pred. No. 65;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQWFWL 10  
| | | | |

Db 545 PKQWGWGL 553

## RESULT 14

Q9PYP2 PRELIMINARY; PRT; 898 AA.  
AC Q9PYP2;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE ORF154.  
GN ORF154.  
OS Xestia c-nigrum granulosis virus (XnGV) (Xestia c-nigrum granulovirus).  
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.  
OX NCBI\_TaxID=51677;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99434230; PubMed=10502508;  
RA Hayakawa T., Ko R., Okano K., Seong S.I., Goto C., Maeda S.;  
RT "Sequence analysis of the Xestia c-nigrum granulovirus genome."; Virology 262:277-297(1999).  
RL EMBL; AF162221; AAF05268.1; -.  
DR EMBL; AF162221; AAF05268.1; -.  
SQ SEQUENCE 898 AA; 104261 MW; DDE9900AEE146834 CRC64;

Query Match 60.6%; Score 43; DB 12; Length 898;  
Best Local Similarity 70.0%; Pred. No. 71; Mismatches 0; Indels 3; Gaps 0;

Matches 7; Conservative 0; Mismatches 0; Indels 3; Gaps 0;

QY 2 PKPQWFWL 11  
| | | | |

Db 353 PYQIWAFLM 362

## RESULT 15

Q34889 PRELIMINARY; PRT; 158 AA.  
AC Q34889;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE HYPOTHETICAL 18.7 KDA PROTEIN YVAV (ORFA).  
GN YVAV.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JH642;  
RA Nakamura A., Grau R., Perego M., Hoch J.A.;  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.

[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borris R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferreri E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Guisepi G., Guy B.J., Haga K., Haech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauei C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Visari A., Wambutt R., Wedler H., Wiedler H., Weltzenegger T.,  
RA Winters K., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the gram-positive bacterium Bacillus  
RT subtilis.";  
RL Nature 390:249-256(1997).  
DR EMBL; AB006738; BAA21900.1; -;  
DR EMBL; 299121; CAB15380.1; -;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 158 AA; 18650 MW; C495D20D39CF46C7 CRC64;

Query Match 57.7%; Score 41; DB 2; Length 158;  
Best Local Similarity 50.0%; Pred. NO. 27;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 RPKQOWFWL 10  
DB 131 ROKPLSYWI 140

RESULT 16  
OB4560 PRELIMINARY; PRT; 159 AA.  
ID OB4560;  
AC OB4560;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DE 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE HYPOTHETICAL 17.8 KDA PROTEIN.  
GN CT556.  
OS Chlamydia trachomatis.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=813;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=D/OW-3/CX;  
RX MEDLINE=99000809; PubMed=9784136;  
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,  
RA Mitchell W.P., Olinger L., Tatusov R.L., Zhao Q., Koonin E.V.,  
RA Davis R.W.;  
RT "Genome sequence of an obligate intracellular pathogen of humans:  
RT Chlamydia trachomatis.";  
RL Science 282:754-759(1998).  
DR EMBL; AE001326; AAC68158.1; -;

KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 159 AA; 17755 MW; 16CC06BAECAF6CBB CRC64;

Query Match 57.7%; Score 41; DB 2; Length 159;  
Best Local Similarity 66.7%; Pred. NO. 27;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 PKPOQWFWL 10  
DB 89 PAPSOWDWL 97

RESULT 17  
Q9DAS6 PRELIMINARY; PRT; 209 AA.  
ID Q9DAS6;  
AC Q9DAS6;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE ADULT FEMALE PLACENTA CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
DE CLONE:1600029015, FULL INSERT SEQUENCE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=PLACENTA;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,  
RA Saito T., Okazaki Y., Gojorori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boifelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK005564; BAB24124.1; -;  
DR InterPro; IPR001063; Ribosomal\_L22.  
DR Pfam; PF00237; Ribosomal\_L22; 1.  
DR ProDom; PD001032; Ribosomal\_L22; 1.  
DR PROSITE; PS00464; RIBOSOMAL\_L22; 1.  
SQ SEQUENCE 209 AA; 24121 MW; AEFEDB4286CD0C3 CRC64;

Query Match 57.7%; Score 41; DB 11; Length 209;  
Best Local Similarity 62.5%; Pred. NO. 35;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PKPOQWFW 9  
DB 66 PRPKQWGW 73

RESULT 18  
Q9HXA6 PRELIMINARY; PRT; 261 AA.  
ID Q9HXA6;  
AC Q9HXA6;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

```

DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL PROTEIN PA3907.
GN PA3907.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01.
RX MEDLINE=20437337; PubMed=10984043;
RA Hickey C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
RW EMBL: AE004808; AAG07294.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 261 AA; 29879 MW; 65556F1E9C330449 CRC64;

Query Match 57.7%; Score 41; DB 2; Length 261;
Best Local Similarity 62.5%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQWFN 9
Db 126 PFPHEWFN 133
| | : | | |

RESULT 19
QSPAVO PRELIMINARY; PRT; 286 AA.
ID QSPAVO;
AC QSPAVO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ACETYLXYLAN ESTERASE.
GN XF2395.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvarenga R., Alves L.M.C., Arya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frome M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Marzocca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Paixoto B.B., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quagdo R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,

```

---

```

RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159 (2000).
DR EMBL: AE004048; AAF85194.1; -.
DR InterPro: IPR002509; Polysac_deacet.
DR Pfam: PF01522; Polysac_deacet; 1.
KW Complete proteome.
SQ SEQUENCE 286 AA; 32166 MW; 013067BBEBC173F CRC64;

Query Match 57.7%; Score 41; DB 2; Length 286;
Best Local Similarity 62.5%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQQWFWM 11
Db 165 PQQWFWM 172
| | : | | |

RESULT 20
Q9AKK5 PRELIMINARY; PRT; 290 AA.
ID Q9AKK5;
AC Q9AKK5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HTRB PROTEIN.
GN HTRB.
OS Rickettsia montana.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsiae; Rickettsia.
OX NCBI_TaxID=33991;
RN [1]
RP SEQUENCE FROM N.A.
RA Andersson J.O., Andersson S.G.E.;
RT "Pseudogenes, junk DNA and the Dynamics of Rickettsia genomes.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ293330; CAC33650.1; -.
SQ SEQUENCE 290 AA; 33635 MW; 9599E3C0C3C076F9 CRC64;

Query Match 57.7%; Score 41; DB 2; Length 290;
Best Local Similarity 55.6%; Pred. No. 48;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQWFN 9
Db 275 KQNPQWFN 283
| | : | | |

RESULT 21
Q55507 PRELIMINARY; PRT; 291 AA.
ID Q55507;
AC Q55507;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE AMMONIUM TRANSPORT PROTEIN.
GN CYSQ OR SLL0895.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RA Tabata S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;

```

RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
RA Suglura M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
RL region from map positions 64% to 92% of the genome.";  
RL DNA Res. 2:153-166(1995).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hiroseawa M., Suglura M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL; D64006; BAA10862.1; -;  
DR InterPro; IPR000760; Inositol\_P.  
DR Pfam; PF00459; Inositol\_P; 1.  
DR PROSITE; PS00630; IMP\_2; 1.  
KW Complete proteome.  
SQ SEQUENCE 291 AA; 32197 MW; F045980AA033E0E3 CRC64;

Query Match 57.7%; Score 41; DB 2; Length 291;  
Best Local Similarity 50.0%; Pred. No. 48;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RPKPQQWFWM 11  
DB 88 PLPQDQWVII 97

RESULT 22

ID Q9HNN1 PRELIMINARY; PRT; 304 AA.  
AC Q9HNN1;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE ACETYLTRANSFERASE HOMOLOG.  
GN YVAI OR VNG2025G.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
OC Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,  
RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leitthaus B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of Halobacterium species NRC-1";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL; AE005096; AAG20189.1; -;  
DR InterPro; IPR001451; Hexapep\_transf.  
DR Pfam; PF00132; hexapep; 4.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 304 AA; 32960 MW; 89EFA2E9A32A921 CRC64;

Query Match 57.7%; Score 41; DB 1; Length 304;  
Best Local Similarity 56.7%; Pred. No. 51;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 9

Db 96 RSKPLOWLW 104  
RESULT 23  
Q9F849 PRELIMINARY; PRT; 392 AA.  
ID Q9F849;  
AC Q9F849;  
DT 01-MAR-2001 (TRENBLrel. 16, Created)  
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)  
DE CMIG.  
OS Streptomyces venezuelae.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=54571;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ISP5230;  
RA He J., Magarvey M.A., Pirae M., Vining L.C.;  
RT "Chloramphenicol Biosynthesis in Streptomyces venezuelae ISP5230:  
RT Functions of Genes Upstream of pabAB";  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF262220; AAG21974.1; -;  
SQ SEQUENCE 392 AA; 43167 MW; 75213EA4FA432CC4 CRC64;

Query Match 57.7%; Score 41; DB 2; Length 392;  
Best Local Similarity 50.0%; Pred. No. 65;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 10  
DB 326 RPEPDRWRW 335

RESULT 24

ID P74332 PRELIMINARY; PRT; 529 AA.  
AC P74332;  
DT 01-FEB-1997 (TRENBLrel. 02, Created)  
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE HYPOTHETICAL 58.0 KDA PROTEIN.  
GN SLR0959.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hiroseawa M., Suglura M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL; D90914; BAA18426.1; -;  
DR InterPro; IPR003675; Abi.  
DR Pfam; PF02517; Abi; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 529 AA; 57992 MW; D2A1B702784AC6B7 CRC64;

Query Match 57.7%; Score 41; DB 2; Length 529;  
Best Local Similarity 70.0%; Pred. No. 87;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 10  
DB 326 RPEPDRWRW 335

```
Db 368 RPLPQDQWFTL 377

RESULT 25
Q9V6W6 PRELIMINARY; PRT; 45 AA.
AC Q9V6W6;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DE CG13352 PROTEIN.
GN CG13352.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtova K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler G., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003816; AAF58301.1; -.
DR FlyBase; FBgn0033894; CG13352.
SQ SEQUENCE 45 AA; 5134 MW; AAAAA100A982418DC CRC64;

Query Match 56.3%; Score 40; DB 5; Length 45;
Best Local Similarity 62.5%; Pred. No. 11;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PKPQQWF 9
1 1 1 1 1
Db 14 PSPQWQW 21
```

```
RESULT 26
Q9R129 PRELIMINARY; PRT; 154 AA.
AC Q9R129;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE RIBONUCLEASE 8 PRECURSOR (FRAGMENT).
GN R8.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Singhanian N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
RA Domachowski J.B., Rosenberg H.F.;
RT "Rapid evolution of the Ribonuclease A superfamily: adaptive expansion
of independent gene clusters in rats and mice.";
RL J. Mol. Evol. 0:0-0(1999).
DR EMBL; AF171646; AAD51666.1; -.
DR HSSP; P00656; ILSO.
DR InterPro; IPR001427; RNaseA.
DR Pfam; PF00074; rnsaeA; 1.
DR PRINTS; PR00794; RIBONUCLEASE.
DR ProDom; PD000535; RNaseA; 1.
DR SMART; SM00092; RNase_Pc; 1.
KW SIGNAL.
FT SIGNAL. 1 25 POTENTIAL.
FT CHAIN 26 >154 RIBONUCLEASE 8.
FT NON_TER 154 154
SQ SEQUENCE 154 AA; 17215 MW; F841AF2931FB2B67 CRC64;

Query Match 56.3%; Score 40; DB 11; Length 154;
Best Local Similarity 75.0%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQWF 8
1 1 1 1 1
Db 27 RPTPSQWF 34

RESULT 27
Q9R132 PRELIMINARY; PRT; 155 AA.
AC Q9R132;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE RIBONUCLEASE 4 PRECURSOR (FRAGMENT).
GN R4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Singhanian N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
RA Domachowski J.B., Rosenberg H.F.;
RT "Rapid evolution of the Ribonuclease A superfamily: adaptive expansion
of independent gene clusters in rats and mice.";
RL J. Mol. Evol. 0:0-0(1999).
DR EMBL; AF171643; AAD51663.1; -.
DR HSSP; P00656; 1RBD.
DR InterPro; IPR001427; RNaseA.
DR Pfam; PF00074; rnsaeA; 1.
DR PRINTS; PR00794; RIBONUCLEASE.
DR ProDom; PD000535; RNaseA; 1.
DR SMART; SM00092; RNase_Pc; 1.
DR PROSITE; PS00127; RNASE_PANCREATIC; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL. 1 25 POTENTIAL.
```



```
FT CHAIN 26 >155 RIBONUCLEASE 4.
FT NON_TER 155 155
SQ SEQUENCE 155 AA; 17223 MW; 5FDF6ECE0A15263C CRC64;

Query Match 56.3%; Score 40; DB 11; Length 155;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPOQWF 8
Db 27 RPTPSQWF 34

RESULT 28
Q9R131 PRELIMINARY; PRT; 155 AA.
AC Q9R131;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE RIBONUCLEASE 5 PRECURSOR (FRAGMENT).
GN R5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Singhanian N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
RA Domachowski J.B., Rosenberg H.F.;
RT "Rapid evolution of the Ribonuclease A superfamily: adaptive expansion
of independent gene clusters in rats and mice.";
RL J. Mol. Evol. 0:0-0(1999).
DR EMBL; AF171644; AAD51664.1; -.
DR HSSP; P00656; 1RBD.
DR InterPro; IPR001427; RNaseA.
DR PRINTS; PR00794; RIBONUCLEASE.
DR PRODOM; PD000535; RNaseA; 1.
DR SMART; SM00092; RNase_Pc; 1.
DR PROSITE; PS00127; RNASE_PANCREATIC; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 >155 RIBONUCLEASE 14.
FT NON_TER 155 155
SQ SEQUENCE 155 AA; 17332 MW; 4F7B83380AA0C6A8 CRC64;

Query Match 56.3%; Score 40; DB 11; Length 155;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPOQWF 8
Db 27 RPTPSQWF 34

RESULT 30
Q9JKI6 PRELIMINARY; PRT; 155 AA.
AC Q9JKI6;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE EOSINOPHIL-ASSOCIATED RIBONUCLEASE 6 PRECURSOR.
GN EAR6.
OS Mus saxicola (Spiny mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10094;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20243759; PubMed=10758160;
RA Zhang J., Dyer K.D., Rosenberg H.F.;
RT "Evolution of the rodent eosinophil-associated ribonuclease gene
family by rapid gene sorting and positive selection.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:4701-4706(2000).
DR EMBL; AF238397; AAF67697.1; -.
DR InterPro; IPR001427; RNaseA.
DR Pfam; PF00074; rnasea; 1.
DR PRINTS; PR00794; RIBONUCLEASE.
DR PRODOM; PD000535; RNaseA; 1.
DR SMART; SM00092; RNase_Pc; 1.
DR PROSITE; PS00127; RNASE_PANCREATIC; UNKNOWN_1.
SQ SEQUENCE 155 AA; 17363 MW; 8B5401F0C26CA1EF CRC64;

Query Match 56.3%; Score 40; DB 11; Length 155;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPOQWF 8
Db 27 RPTPSQWF 34

RESULT 31
Q9KRQ9 PRELIMINARY; PRT; 273 AA.
AC Q9KRQ9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
```

```
FT CHAIN 26 >155 RIBONUCLEASE 4.
FT NON_TER 155 155
SQ SEQUENCE 155 AA; 17223 MW; 5FDF6ECE0A15263C CRC64;

Query Match 56.3%; Score 40; DB 11; Length 155;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPOQWF 8
Db 27 RPTPSQWF 34

RESULT 28
Q9R131 PRELIMINARY; PRT; 155 AA.
AC Q9R131;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE RIBONUCLEASE 5 PRECURSOR (FRAGMENT).
GN R5.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Singhanian N.A., Dyer K.D., Zhang J., Deming M.S., Bonville C.A.,
RA Domachowski J.B., Rosenberg H.F.;
RT "Rapid evolution of the Ribonuclease A superfamily: adaptive expansion
of independent gene clusters in rats and mice.";
RL J. Mol. Evol. 0:0-0(1999).
DR EMBL; AF171644; AAD51664.1; -.
DR HSSP; P00656; 1RBD.
DR InterPro; IPR001427; RNaseA.
DR Pfam; PF00074; rnasea; 1.
DR PRINTS; PR00794; RIBONUCLEASE.
DR PRODOM; PD000535; RNaseA; 1.
DR SMART; SM00092; RNase_Pc; 1.
DR PROSITE; PS00127; RNASE_PANCREATIC; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 >155 RIBONUCLEASE 5.
FT NON_TER 155 155
SQ SEQUENCE 155 AA; 17189 MW; 36C6F3381DB92787 CRC64;

Query Match 56.3%; Score 40; DB 11; Length 155;
Best Local Similarity 75.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPOQWF 8
Db 27 RPTPSQWF 34

RESULT 29
Q9R127 PRELIMINARY; PRT; 155 AA.
AC Q9R127;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE RIBONUCLEASE 14 PRECURSOR (FRAGMENT).
GN R14.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HYPOTHETICAL PROTEIN VC1577.
GN VC1577.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL: AE004235; RAF94731.1; -.
DR TIGR: VC1577; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 273 AA; 31087 MW; 9A21040F4DBA773E CRC64;

Query Match 56.3%; Score 40; DB 2; Length 273;
Best Local Similarity 71.4%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQFWL 10
Db 262 PEQIWL 268
|:|:|

RESULT 32
Q9JYV2 PRELIMINARY; PRT; 289 AA.
AC Q9JYV2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE HTRB/MSBB FAMILY PROTEIN.
GN NMB1418.
OS Neisseria meningitidis (serogroup B).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=491;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MC58 / SEROGROUP B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Clecko A., Parksey D.S., Blair E., Citterio H., Clark E.B.,
RA Cotton M.D., Uitterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlato V., Maignani V., Pizza M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58."
RL Science 287:1809-1815(2000).
DR EMBL: AE002491; AAP41779.1; -.
DR TIGR: NMB1418; -.
KW Complete proteome.
SQ SEQUENCE 289 AA; 33843 MW; 3BF100F050576512 CRC64;

Query Match 56.3%; Score 40; DB 2; Length 289;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 1 RPKPQQFWL 10
Db 265 REHQYFWL 274
|:|:|

RESULT 33
Q9JUT4 PRELIMINARY; PRT; 289 AA.
AC Q9JUT4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE PUTATIVE ACETYLTRANSFERASE.
GN NMA1630.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies K.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
RA Whitehead S., Spratt B.G., Barrall B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RL Nature 404:502-506(2000).
DR EMBL: ALI62756; CAB84858.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 289 AA; 33867 MW; 4FAE453B5A632C1D CRC64;

Query Match 56.3%; Score 40; DB 2; Length 289;
Best Local Similarity 60.0%; Pred. No. 69;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10
Db 265 REHQYFWL 274
|:|:|

RESULT 34
Q9ZCL1 PRELIMINARY; PRT; 290 AA.
AC Q9ZCL1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE (HTRB).
GN RP718.
OS Rickettsia prowazekii.
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
OC Rickettsiaceae; Rickettsia.
OX NCBI_TaxID=782;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MADRID E;
RX MEDLINE=99039499; PubMed=9823893;
RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA Scharitz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT "The genome sequence of Rickettsia prowazekii and the origin of
RT mitochondria."
RL Nature 396:133-140(1998).
DR EMBL: AJ235273; CAA15149.1; -.
KW Complete proteome.
SQ SEQUENCE 290 AA; 33809 MW; 991FF50AB841D5B3 CRC64;

Query Match 56.3%; Score 40; DB 2; Length 290;
```

Best Local Similarity 55.6%; Pred. No. 69;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPQWF 9

Db 275 KONPAQWF 283

RESULT 35

Q9AKF0  
ID Q9AKF0 PRELIMINARY; PRT; 290 AA.  
AC Q9AKF0;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE HTRB PROTEIN.  
GN HTRB.  
OS Rickettsia rickettsii.  
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
OC Rickettsiaceae; Rickettsiae; Rickettsia.  
OX NCBI\_TaxID=783;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=84-21C;  
RA Andersson J.O., Andersson S.G.E.;  
RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes."  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ293329; CAC33715.1; -.  
SQ SEQUENCE 290 AA; 33550 MW; 9FC5D73E5CBE89A CRC64;

Query Match 56.3%; Score 40; DB 2; Length 290;  
Best Local Similarity 55.6%; Pred. No. 69;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPQWF 9

Db 275 KONPAQWF 283

RESULT 36

Q9AKA6  
ID Q9AKA6 PRELIMINARY; PRT; 290 AA.  
AC Q9AKA6;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE HTRB PROTEIN.  
GN HTRB.  
OS Rickettsia typhi.  
OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;  
OC Rickettsiaceae; Rickettsiae; Rickettsia.  
OX NCBI\_TaxID=785;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WILMINGTON;  
RA Andersson J.O., Andersson S.G.E.;  
RT "Pseudogenes, junk DNA and the dynamics of Rickettsia genomes."  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ293328; CAC33762.1; -.  
SQ SEQUENCE 290 AA; 33814 MW; 2E9B2D3193B70EDD CRC64;

Query Match 56.3%; Score 40; DB 2; Length 290;  
Best Local Similarity 55.6%; Pred. No. 69;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPQWF 9

Db 275 KONPAQWF 283

RESULT 37

Q9KVD3  
ID Q9KVD3 PRELIMINARY; PRT; 318 AA.  
AC Q9KVD3;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE.  
GN VC0213.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
EX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickley E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischman R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.W.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
cholerae."  
RL Nature 406:477-483(2000).  
DR EMBL; AE004111; AAF93389.1; -.  
RW TIGR; VC0213; -.  
KW Transferase; Acyltransferase; Complete proteome.  
SQ SEQUENCE 318 AA; 36542 MW; FE95D7A4C83106E1 CRC64;

Query Match 56.3%; Score 40; DB 2; Length 318;  
Best Local Similarity 71.4%; Pred. No. 76;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PQQFWL 10

Db 294 PQQFWL 300

RESULT 38

O82937  
ID O82937 PRELIMINARY; PRT; 332 AA.  
AC O82937;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE LIPID A BIOSYNTHESIS (KDO)2-(LAUROYL)-LIPID IVA ACYLTRANSFERASE.  
GN ECF4 OR MSBB  
OS Escherichia coli O157:H7.  
OC Plasmid pO157.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98290540; PubMed=9628576;  
RA Makino K., Ishii K., Yasunaga T., Hattori M., Yokoyama K.,  
RA Yatsudo H.C., Kubota Y., Yamauchi Y., Iida T., Yamamoto K., Honda T.,  
RA Han C., Ohtsubo A., Kasamatsu M., Hayashi T., Kuhara S., Shinagawa H.,  
RT "Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an  
enterohemorrhagic Escherichia coli O157:H7 derived from Sakai  
outbreak."  
RT DNA Res. 5:1-9(1998).  
RN [2]  
RP SEQUENCE OF 1-235 FROM N.A.  
RC STRAIN=4304-PT14;  
RX MEDLINE=98261495; PubMed=9596716;  
RA Boerlin P., Chen S., Colbourne J.K., Johnson R., De Grandis S.,  
RA Gyles C.;  
RT "Evolution of enterohemorrhagic Escherichia coli hemolysin plasmids  
and the locus for enterocyte effacement in shiga toxin-producing E.  
coli."



Q9FYI5;  
AC 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DT F17F8.14.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Khan S., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Kim C.,  
RA Shinn P., Altafi H., Bei O., Chin C., Chiou J., Choi E., Conn L.,  
RA Conway A., Gonzales A., Hansen N., Howng B., Koo T., Lam B., Lee J.,  
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
RA Ecker J.R.;  
RT "Genomic sequence for Arabidopsis thaliana BAC F17F8 from chromosome  
I.,";  
RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RA Shinn P., Brooks S., Buehler E., Chao Q., Dunn P., Khan S., Kim C.,  
RA Walker M., Altafi H., Araujo R., Conn L., Conway A., Gonzalez A.,  
RA Hansen N., Huizar L., Kremetskaia I., Lenz C., Li J., Liu S.,  
RA Luros S., Rowley D., Schwartz J., Toriumi M., Vysotskaia V., Yu G.,  
RA Davis R., Federspiel N., Theologis A., Ecker J.;  
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,  
RA Conn L., Conway A., Gonzalez A., Hansen N., Howng B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
RA Theologis A., Ecker J.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [7]  
DR EMBL; AC000107; AAF98186.1; -;  
DR InterPro; IPR003656; BED\_finger.  
DR InterPro; IPR002965; P\_rich\_extensn.  
DR InterPro; IPR000822; Znf-C2H2.  
DR Pfam; PF00096; zf-C2H2; 2.  
DR PRINTS; PR01217; PRICHEXTENS.  
DR SMART; SM00355; Znf\_C2H2; 2.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; 1.  
KW DNA-binding; Metal-binding; Zinc-finger.  
SQ SEQUENCE 381 AA; 41208 MW; ACE3B529D0022154 CRC64;

Query Match 56.3%; Score 40; DB 10; Length 381;  
Best Local Similarity 71.4%; Pred. No. 90;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 PKPQQWF 8  
DB 161 PRPQQWF 167

RESULT 43  
Q9NRT2  
AC Q9NRT2 PRELIMINARY; PRT; 411 AA.  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE TWO-PORE DOMAIN POTASSIUM CHANNEL TREK-1.  
GN TREK-1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Meadows H.J., Benham C.D., Cairns W., Gloger I.S., Jennings C.,  
RA Medhurst A.D., Murdock P., Chapman C.G.;  
RT "Cloning, localization and functional expression of the human ortholog  
of the TREK-1 potassium channel.,";  
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF171068; AAF89743.1; -;  
DR InterPro; IPR003280; 2porek\_channel.  
DR InterPro; IPR001622; Channel\_pore\_K.  
DR InterPro; IPR003976; Trek\_channel.  
DR PRINTS; PR01333; 2POREKCHANNEL.  
DR PRINTS; PR01499; TREKCHANNEL.  
KW Ionic channel.  
SQ SEQUENCE 411 AA; 45494 MW; FDE40CAB21B42A1C CRC64;  
Query Match 56.3%; Score 40; DB 4; Length 411;  
Best Local Similarity 55.6%; Pred. No. 97;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 3 KPQQWFWM 11  
DB 271 KPQQWFWM 279  
RESULT 44  
Q9H8K3  
AC Q9H8K3 PRELIMINARY; PRT; 639 AA.  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE CDNA FLJ13515 FIS, CLONE PLACE1005595.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PLACENTA;  
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,  
RA Nakamura H., Ishii S., Kawai Y., Salto K., Yamamoto J., Wakamatsu A.,  
RA Nakamura Y., Nagahari K., Masuho Y., Sasaki N.;  
RT "NEO human cDNA sequencing project.,";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK023577; BAB14613.1; -;  
SQ SEQUENCE 639 AA; 68405 MW; 739FF6209738D832 CRC64;  
Query Match 56.3%; Score 40; DB 4; Length 639;  
Best Local Similarity 55.6%; Pred. No. 1.5e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 2 KPQQWFWM 10  
DB 506 POPHNWVL 514

```
RESULT 45
Q9V7L5 ID Q9V7L5 PRELIMINARY; PRT; 665 AA.
AC Q9V7L5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CNG PROTEIN.
GN CNG OR CG7779.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.A., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000)
DR EMBL; AE003807; AAF58033.1;
DR FlyBase; FBgn0014462; Cng.
DR InterPro; IPR001622; Channel_pore_K.
DR InterPro; IPR002025; Cng_membrane.
DR InterPro; IPR000595; cNMP_binding.
DR Pfam; PF00914; CNG_membrane; 1.
DR Pfam; PF00027; cNMP_binding; 1.
DR SMART; SM00100; cNMP; 1.
DR PROSITE; PS00888; cNMP_BINDING_1; 1.
DR PROSITE; PS00889; cNMP_BINDING_2; 1.
DR PROSITE; PS50042; cNMP_BINDING_3; 1.
SQ SEQUENCE 665 AA; 75823 MW; 6EFFC9A7CA243660 CRC64;
```

Query Match 56.3%; Score 40; DB 5; Length 665;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;

```
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 RPKPOQWF 8
   |||||
Db 52 RPKPPDWF 59

RESULT 46
Q9NWYO ID Q9NWYO PRELIMINARY; PRT; 774 AA.
AC Q9NWYO;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE CDNA FLJ20539 FIS, CLONE KAT11311.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,
RA Suzuki Y., Ohtsubashi M., Nishitani T., Shibahara T., Tanaka T.,
RA Nakamura Y., Isono T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000546; BAA91245.1;
SQ SEQUENCE 774 AA; 83181 MW; 8C09E73DF939A7ED CRC64;
```

Query Match 56.3%; Score 40; DB 4; Length 774;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
QY 2 PKPOQFWL 10
   |||
Db 641 PQPHNWVL 649
```

```
RESULT 47
ID Q9CAX2 PRELIMINARY; PRT; 797 AA.
AC Q9CAX2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOPHETICAL 89.4 KDA PROTEIN.
GN F24K9.23.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansong W., Unselid M.,
RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choise N., Artiguenave F., Robert C., Brottier P.,
RA Wincker P., Cattolico L., Weissbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erffle H., Jordan N., Bangert S.,
RA Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppi S., Simonati B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordsiek G.,
RA Reichelt J., Scharfe M., Schoen O., Bagues M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy B.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
```

```

RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Miltscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RA "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
DR ENBL; AC008153; AAG51444.1; -.
KW Hypothetical protein.
SQ SEQUENCE 797 AA; 89393 MW; 2E68459765D08B52 CRC64;

Query Match 56.3%; Score 40; DB 10; Length 797;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PKPQOWFW 9
Db 188 PKAYEWF 195

RESULT 48
ID Q9HE57 PRELIMINARY; PRT; 937 AA.
AC Q9HE57;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE RELATED TO SLS2 PROTEIN.
GN B2F7.50.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,
RA Nyakatura G., Mewes H.W., Mannhaupt G.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL451013; CAC18157.1; -.
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001230; Prenyltn.
DR Pfam; PF00646; F-box; 1.
DR SMART; SM00256; FBOX; 1.
DR PROSITE; PS00181; FBOX; 1.
DR PROSITE; PS00294; PRENYLATIN; UNKNOWN 1.
SQ SEQUENCE 937 AA; 104264 MW; 2F6D7A68FA851FA7 CRC64;

Query Match 56.3%; Score 40; DB 3; Length 937;
Best Local Similarity 60.0%; Pred. No. 2.2e+02;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PKPQOWFW 11
Db 349 PKPAEWF 358

RESULT 49
Q9RV51 PRELIMINARY; PRT; 548 AA.
ID Q9RV51
AC Q9RV51;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

```

```

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL 59.4 KDA PROTEIN.
GN DR1179.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI.
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Mofat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Manton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001967; AAF10758.1; -.
DR TIGR; DR1179; -.
DR InterPro; IPR001736; PLD.
DR Pfam; PF00614; PLDC; 2.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 548 AA; 59442 MW; AEA550DBEE981B0 CRC64;

Query Match 55.6%; Score 39.5; DB 2; Length 548;
Best Local Similarity 70.0%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 RPKPOQFW 10
Db 521 RVPQEW-WL 529

RESULT 50
O79563 PRELIMINARY; PRT; 55 AA.
ID O79563
AC O79563;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ATPASE SUBUNIT 8.
GN ATP8.
OS Mustelus manazo.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes; Triakidae;
OC Mustelus.
OX NCBI_TaxID=79736;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER.
RA Cao Y., Waddell P.J., Okada N., Hasegawa M.;
RT "The complete mitochondrial DNA sequence of the shark (Mustelus
RT manazo): Evaluating rooting contradictions to living bony
RT vertebrates.";
RL Mol. Biol. Evol. 0:0-0(1998).
DR EMBL; AB015962; BAA33040.1; -.
DR InterPro; IPR001421; ATP-synt_8.
DR Pfam; PF00895; ATP-synt_8; 1.
KW Mitochondrion.
SQ SEQUENCE 55 AA; 6616 MW; 52DB099625A0D75D CRC64;

Query Match 54.9%; Score 39; DB 8; Length 55;
Best Local Similarity 55.6%; Pred. No. 20;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPOQFW 9

```

```
DB 44 KPKPNPNW 52
RESULT 51
Q69572 PRELIMINARY; PRT; 99 AA.
AC Q69572;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-NOV-1998 (TEMBLrel. 08, Last annotation update)
DE U96 PROTEIN.
GN U96 OR HCLF1.
OS Human herpesvirus 6.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Roseolovirus.
OX NCBI_TaxID=10368;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=95266321; PubMed=7747482;
RA Compels U.A., Nicholas J., Lawrence G., Jones M., Thomson B.J.,
RA Martin M.E., Efstathiou S., Craxton M., Macaulay H.A.;
RT "The DNA sequence of human herpesvirus-6: structure, coding content,
RT and genome evolution."
RL J. Virol. 209:29-51(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=90080132; PubMed=2152817;
RA Lawrence G.L., Chee M., Craxton M.A., Compels U.A., Honess R.W.,
RA Barrell B.G.;
RT "Human herpesvirus 6 is closely related to human cytomegalovirus."
RL J. Virol. 64:287-299(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=91237802; PubMed=1851860;
RA Chang C.K., Balachandran N.;
RT "Identifying a phosphoprotein of human herpesvirus 6."
RL J. Virol. 65:2884-2894(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=91333007; PubMed=1651403;
RA Teo I.A., Griffin B.E., Jones M.D.;
RT "Characterization of the DNA polymerase gene of human herpesvirus 6."
RL J. Virol. 65:4670-4680(1991).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=91226542; PubMed=1851252;
RA Thomson B.J., Efstathiou S., Honess R.W.;
RT "Acquisition of the human adeno-associated virus type-2 rep gene by
RT human herpesvirus type-6."
RL Nature 351:78-80(1991).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=91374590; PubMed=1654446;
RA Martin M.E., Nicholas J., Thomson B.J., Newman C., Honess R.W.;
RT "Identification of a transactivating function mapping to the putative
RT immediate-early locus of human herpesvirus 6."
RL J. Virol. 65:5381-5390(1991).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=92333249; PubMed=1321206;
RA Efstathiou S., Lawrence G.L., Brown C.M., Barrell B.G.;
RT "Identification of homologues to the human cytomegalovirus US22 gene
RT family in human herpesvirus 6."
RL J. Gen. Virol. 73:1661-1671(1992).
RN [8]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=92148942; PubMed=1310766;
RA Geng Y., Chandran B., Josephs S.F., Wood C.;
RT "Identification and characterization of a human herpesvirus 6 gene
RT segment that trans activates the human immunodeficiency virus type 1
RT promoter."
RL J. Virol. 66:1564-1570(1992).
RN [9]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=93091236; PubMed=1333836;
RA Compels U.A., Carss A.L., Sun N., Arrand J.R.;
RT "Infectivity determinants encoded in a conserved gene block of human
RT herpesvirus-6."
RL DNA Seq. 3:25-39(1992).
RN [10]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=92260671; PubMed=1374813;
RA Neipel F., Ellinger K., Fleckenstein B.;
RT "Gene for the major antigenic structural protein (p100) of human
RT herpesvirus 6."
RL J. Virol. 66:3918-3924(1992).
RN [11]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=92333248; PubMed=1321205;
RA Thomson B.J., Honess R.W.;
RT "The right end of the unique region of the genome of human herpesvirus
RT 6 U1102 contains a candidate immediate early gene enhancer and a
RT homologue of the human cytomegalovirus US22 gene family."
RL J. Gen. Virol. 73:1649-1660(1992).
RN [12]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=93187613; PubMed=8383182;
RA Ellinger K., Neipel F., Foa-Tomasi L., Campadelli-Fiume G.,
RA Fleckenstein B.;
RT "The glycoprotein B homologue of human herpesvirus 6."
RL J. Gen. Virol. 74:495-500(1993).
RN [13]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=93224882; PubMed=8385692;
RA Compels U.A., Carrigan D.R., Carss A.L., Arno J.;
RT "Two groups of human herpesvirus 6 identified by sequence analyses of
RT laboratory strains and variants from Hodgkin's lymphoma and bone
RT marrow transplant patients."
RL J. Gen. Virol. 74:613-622(1993).
RN [14]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=93389439; PubMed=8397282;
RA Liu D.X., Compels U.A., Nicholas J., Lelliott C.;
RT "Identification and expression of the human herpesvirus 6 glycoprotein
RT H and interaction with an accessory 40K glycoprotein."
RL J. Gen. Virol. 74:1847-1857(1993).
RN [15]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=94025558; PubMed=7692666;
RA Liu D.X., Compels U.A., Foa-Tomasi L., Campadelli-Fiume G.;
RT "Human herpesvirus-6 glycoprotein H and L homologs are components of
RT the gp100 complex and the gH external domain is the target for
RT neutralizing monoclonal antibodies."
RL Virology 197:12-22(1993).
RN [16]
RP SEQUENCE FROM N.A.
RC STRAIN=U1102, VARIANT A;
RX MEDLINE=93331710; PubMed=7687803;
RA Pellett P., Sanchez-Martinez D., Dominguez G., Black J.B., Anton E.,
```



```
RA Greenamoyer C., Dambaugh T.R.;
RT "A strongly immunoreactive virion protein of human herpesvirus 6
RT variant B strain 229: identification and characterization of the gene
RT and mapping of a variant-specific monoclonal antibody reactive
RT epitope.";
RL Virology 195:521-531(1993).
[17]
RN
RP SEQUENCE FROM N.A.
RX STRAIN-U1102, VARIANT A;
RX MEDLINE-93323202; PubMed=7687301;
RA Pfeiffer B., Berneman Z.N., Neipel F., Chang C.K., Tirwatnapong S.,
RA Chandran B.;
RT "Identification and mapping of the gene encoding the glycoprotein
RT complex gp82-gp105 of human herpesvirus 6 and mapping of the
RT neutralizing epitope recognized by monoclonal antibodies.";
RL J. Virol. 67:4611-4620(1993).
[18]
RN
RP SEQUENCE FROM N.A.
RX STRAIN-U1102, VARIANT A;
RX MEDLINE-95146989; PubMed=7844567;
RA Compels U.A., Macaulay H.A.;
RT "Characterization of human telomeric repeat sequences from human
RT herpesvirus 6 and relationship to replication.";
RL J. Gen. Virol. 76:451-458(1995).
[19]
RN
RP SEQUENCE FROM N.A.
RX STRAIN-U1102, VARIANT A;
RX MEDLINE-94047392; PubMed=8230490;
RA Dewhurst S., Dollard S.C., Pellett P.E., Dambaugh T.R.;
RT "Identification of a lytic-phase origin of DNA replication in human
RT herpesvirus 68 strain 229.";
RL J. Virol. 67:7680-7683(1993).
[20]
RN
RP SEQUENCE FROM N.A.
RX STRAIN-U1102, VARIANT A;
RX Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
[21]
RN
RP SEQUENCE FROM N.A.
RX STRAIN-U1102, VARIANT A;
RX MEDLINE-94118404; PubMed=8289364;
RA Nicholas J., Martin M.;
RT "Nucleotide sequence analysis of a 38.5-kilobase-pair region of the
RT genome of human herpesvirus 6 encoding human cytomegalovirus
RT immediate-early gene homologs and transactivating functions.";
RL J. Virol. 68:597-610(1994).
[22]
RN
RP SEQUENCE FROM N.A.
RX STRAIN-U1102, VARIANT A;
RX MEDLINE-94202284; PubMed=8151768;
RA Schieve U., Neipel F., Schreiner D., Fleckenstein B.;
RT "Structure and transcription of an immediate-early region in the human
RT herpesvirus 6 genome.";
RL J. Virol. 68:2978-2985(1994).
[23]
RN
RP SEQUENCE FROM N.A.
RX STRAIN-U1102, VARIANT A;
RX MEDLINE-94181269; PubMed=8134119;
RA Thompson J., Choudhury S., Kashanchi F., Doniger J., Berneman Z.,
RA Frenkel N., Rosenthal L.J.;
Query Match 54.9%; Score 39; DB 12; Length 99;
Best Local Similarity 50.0%; Pred. No. 35;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 4 PQQWFWM 11
Db 86 PSRWYWL 93
RESULT 52
Q55572
ID Q55572 PRELIMINARY; PRT; 208 AA.
```

```
Q55572;
AC 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 23.3 KDA PROTEIN.
GN SLL0156.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-PCC6803;
RA Tabata S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
[2]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE-96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugiyura M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
[3]
RN
RP SEQUENCE FROM N.A.
RX MEDLINE-97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirose M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
KW EMBL; D63999; BAA10075.1; -.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 208 AA; 23268 MW; 327BE1A8BCBBA0E CRC64;

Query Match 54.9%; Score 39; DB 2; Length 208;
Best Local Similarity 55.6%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
Qy 2 PKPQWFWM 10
Db 34 PSPQPMQWI 42
RESULT 53
QJUSJ3
ID QJUSJ3 PRELIMINARY; PRT; 341 AA.
AC QJUSJ3;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2000 (TReMBLrel. 15, Last annotation update)
DE ACYLTRANSFERASE.
GN HTRB_2.
OS Chlamydia pneumoniae (Chlamydia pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83558;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-J138;
RX MEDLINE-20330349; PubMed=10871362;
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RT from Japan and CWL029 from USA.";
RL Nucleic Acids Res. 28:2311-2314(2000).
DR EMBL; AP002545; BAA98308.1; -.
KW Transferase.
SQ SEQUENCE 341 AA; 39044 MW; ECF7A78E1615896A CRC64;
```

```
Query Match          54.9%; Score 39; DB 2; Length 341;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQWFNL 10
   :|:|:|
Db 180 QPEQMWI 187

RESULT 54
Q9PIW6          PRELIMINARY; PRT; 361 AA.
AC Q9PIW6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ALPHA-(1.3/1.4)-FUCOSYLTRANSFERASE.
GN FUT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP Nishihara S.;
SEQUENCE FROM N.A.
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99190658; PubMed=10089211;
RA Nishihara S., Hiraga T., Ikehara Y., Iwasaki H., Kudo T., Yazawa S.,
RA Morozumi K., Suda Y., Narimatsu H.;
RT "Molecular behavior of mutant Lewis enzymes in vivo.";
RL Glycobiology 9:373-382(1999).
DR EMBL; AB043998; BAA96390.1; -
DR InterPro; IPR001503; Glyco_transf_10.
DR Pfam; PF00852; Glyco_transf_10; 1.
KW Transferase; Glycosyltransferase.
FT VARIANT 20 20 R->L
SQ SEQUENCE 361 AA; 42160 MW; 6157E6B63BC34E5E CRC64;

Query Match          54.9%; Score 39; DB 4; Length 361;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQFWF 9
   ||:|:|
Db 126 RPOGQRIW 134

RESULT 55
Q9F669          PRELIMINARY; PRT; 383 AA.
AC Q9F669
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PILIN BIOGENESIS PROTEIN PILC.
GN PILC.
OS Pseudomonas fluorescens.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=294;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WCS365;
RA Camacho Carvajal M.M., de Priester W., Lugtenberg B.J.J.,
RA Bloemberg G.V.;
RT "Involvement of type 4 pilin of Pseudomonas fluorescens in tomato root
RT colonization.";
RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF297457; AAG18588.1; -.
DR InterPro; IPR001992; Bact_secr_systII.
DR Pfam; PF00482; GSPII_F; 1.
DR PRINTS; PR00812; BCTERIAIGSPF.
DR PROSITE; PS00874; T2SP_F; 1.
SQ SEQUENCE 383 AA; 41752 MW; B0556D4ACD8FCDDF CRC64;

Query Match          54.9%; Score 39; DB 2; Length 333;
Best Local Similarity 71.4%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 QQWFWM 11
   ||:|:|
Db 200 QQWWMVW 206

RESULT 56
Q9ZJ55          PRELIMINARY; PRT; 412 AA.
AC Q9ZJ55
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE UBIQUINOL CYTOCHROME C OXIDOREDUCTASE, CYTOCHROME B SUBUNIT.
GN CYTB OR PETB OR JHP1460.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
CC -1- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -1- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.
DR EMBL; AE001568; AAD07046.1; -.
DR InterPro; IPR000179; Cyt_b_b6.
DR Pfam; PF00032; cytochrome_b_c1; 1.
DR Pfam; PF00033; cytochrome_b_n; 1.
DR PROSITE; PS00193; CYTOCHROME_B_OO; UNKNOWN_1.
KW Complete proteome; Electron transport; Heme; Respiratory chain;
Transmembrane.
SQ SEQUENCE 412 AA; 47631 MW; 363F3CDB3638B0BD CRC64;

Query Match          54.9%; Score 39; DB 2; Length 412;
Best Local Similarity 54.5%; Pred. No. 1.4e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQQFWF 11
   ||:|:|:|
Db 349 RPAFWMVFWLL 359

RESULT 57
Q9HU99          PRELIMINARY; PRT; 416 AA.
ID Q9HU99
```

AC Q9HU99;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE PROBABLE OXIDOREDUCTASE.  
GN PAS084.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-O.T., Erwin A.L., Mizoquchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Salier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
DR EMBL; AE004921; AG08469.1; -.  
DR InterPro; IPR000927; DAO.  
DR InterPro; IPR000205; NAD\_binding.  
DR InterPro; IPR000309; TrkA\_Kuptake.  
DR Pfam; PF01266; DAO; 1.  
DR PRINTS; PR00335; KUPTAKETRA.  
KW Complete proteome.  
SQ SEQUENCE 416 AA; 44767 MW; 1EDE536FD784FF6F CRC64;  
  
Query Match 54.9%; Score 39; DB 2; Length 416;  
Best Local Similarity 53.8%; Pred. No. 1.4e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 2; Gaps 1;  
  
QY 1 RPK--PQOWFWLM 11  
II: I I I I I:  
DB 80 RRLDPAQNRWLL 92  
  
RESULT 58  
Q9PIU6 PRELIMINARY; PRT; 452 AA.  
AC Q9PIU6;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE PUTATIVE INTEGRAL MEMBRANE PROTEIN WITH HAEMOLYSIN DOMAIN.  
GN CJ0183.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
OC Campylobacter.  
OX NCBI\_TaxID=197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NCTC 11168;  
RX MEDLINE=20150912; PubMed=10688204;  
RA Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,  
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,  
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., Van Vliet A.H.M.,  
RA Whitehead S., Barrell B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
RT reveals hypervariable sequences.";  
RL Nature 403:665-668(2000).  
DR EMBL; AL139074; CAB72666.1; -.  
DR InterPro; IPR000644; CBS.  
DR InterPro; IPR002550; DUF21.  
DR Pfam; PF00571; CBS; 2.  
DR Pfam; PF01595; DUF21; 1.  
DR SMART; SM00116; CBS; 2.

KW Complete proteome.  
SQ SEQUENCE 452 AA; 51010 MW; E508BD021CA5C79E CRC64;  
  
Query Match 54.9%; Score 39; DB 2; Length 452;  
Best Local Similarity 44.4%; Pred. No. 1.5e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 3 KPOQWFWLM 11  
II: I I I I I:  
DB 160 RPLHWFWM 168  
  
RESULT 59  
Q9K217 PRELIMINARY; PRT; 462 AA.  
AC Q9K217;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE LIPID A BIOSYNTHESIS LAUROYL ACYLTRANSFERASE, PUTATIVE.  
GN CP0676.  
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.  
OX NCBI\_TaxID=83558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AR39;  
RX MEDLINE=20150255; PubMed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
RA White O., Hickey E.K., Peterson J., Umayam L.A., Utterback T.,  
RA Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B.,  
RA Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J.,  
RA McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
RT pneumoniae AR39.";  
RL Nucleic Acids Res. 28:1397-1406(2000).  
DR EMBL; AE002225; AAF38487.1; -.  
DR TIGR; CP0676; -.  
DR InterPro; IPR000504; RRM.  
DR PROSITE; PS00030; RRM\_RNP\_1; UNKNOWN\_1.  
KW Transferase; Acyltransferase.  
SQ SEQUENCE 462 AA; 52624 MW; C31CD0C9B01E13DE CRC64;  
  
Query Match 54.9%; Score 39; DB 2; Length 462;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 KPOQWFWL 10  
II: I I I I I:  
DB 301 QPEQWMMI 308  
  
RESULT 60  
Q9Z983 PRELIMINARY; PRT; 467 AA.  
AC Q9Z983;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE ACYLTRANSFERASE.  
GN HTRB OR CPN0098.  
OS Chlamydia pneumoniae (Chlamydophila pneumoniae).  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.  
OX NCBI\_TaxID=83558;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CWL029;  
RX MEDLINE=99206606; PubMed=10192388;  
RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,  
RA Ollinger L., Grimwood J., Davis R.W., Stephens R.S.;  
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";

```

RL Nat. Genet. 21:385-389(1999).
DR EMBL; AE001596; AAD18251.1; -.
KW Transferase; Complete proteome.
SQ SEQUENCE 467 AA; 53193 MW; D3C7C284E9220DD0 CRC64;

Query Match 54.9%; Score 39; DB 2; Length 467;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 3 KPQQWFWM 10
   :|:|:|:
Db 301 QPEQWMI 308

RESULT 61
OYVDN6 PRELIMINARY; PRT; 480 AA.
AC OYVDN6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HPOTHEUTICAL 54.4 KDA PROTEIN APE0879.
GN APE0879.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kwarabavasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh A.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000080; BAA79861.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 480 AA; 54363 MW; C8F617CCEFF11627E CRC64;

Query Match 54.9%; Score 39; DB 1; Length 480;
Best Local Similarity 55.6%; Pred. No. 1.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 3 KPQQWFWM 11
   ||:|:|:
Db 283 KPEQWFI 291

RESULT 62
OYVJ81 PRELIMINARY; PRT; 499 AA.
AC OYVJ81;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CG10178 PROTEIN.
OS CG10178.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;

```

```

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bens P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstien P., Bottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Hojck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacלב J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003659; AAF53673.1; -.
DR FlyBase; FBgn0032684; CG10178.
DR InterPro; IPR002213; UDPGT.
DR Pfam; PF00201; UDPGT; 2.
DR PROSITE; PS00375; UDPGT; 1.
SQ SEQUENCE 499 AA; 57924 MW; 3DCD133EC526883A CRC64;

Query Match 54.9%; Score 39; DB 5; Length 499;
Best Local Similarity 55.6%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 2 KPQQWFWM 10
   | | | | |
Db 34 PAPSHWLWL 42

RESULT 63
O9HB16 PRELIMINARY; PRT; 524 AA.
AC O9HB16;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE CYP4F11.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, AND LIVER;
RX MEDLINE=20422669; PubMed=10964514;
RA Cui X., Nelson D.R., Strobel H.W.;
RT "A novel human cytochrome p450 4F isoform (CYP4F11): cDNA cloning,

```

RT expression, and genomic structural characterization.";  
RL Genomics 68:161-166(2000).  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
DR EMBL: AF236085; AAG15889.1; -.  
DR InterPro: IPR001128; Cyt\_P450.  
DR Pfam: PF00067; P450; 1.  
DR PRINTS: PR00385; P450.  
DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
KW Heme; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 524 AA; 60209 MW; 627774C85904B918 CRC64;

Query Match 54.9%; Score 39; DB 4; Length 524;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPQPQWF 9  
: | | | |  
DB 53 QPQKQWF 61

RESULT 64  
O83145 PRELIMINARY; PRT; 525 AA.  
ID O83145;  
AC O83145;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE LICC PROTEIN (LICC).  
GN TP0107.  
OS Treponema pallidum.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
OX NCBI\_TaxID=160;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98332770; PubMed=9665876;  
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,  
RA Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,  
RA Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,  
RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,  
RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
RA Venter J.C.;  
RT "Complete genome sequence of Treponema pallidum, the syphilis  
spirochete.";  
RL Science 281:375-388(1998).  
DR EMBL: AE001195; AAC26555.1; -.  
DR TIGR: TP0107; -.  
DR InterPro: IPR001245; TYF\_kin.  
DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1.  
KW Complete proteome.  
SQ SEQUENCE 525 AA; 61044 MW; E46DDCACC79BAB10 CRC64;

Query Match 54.9%; Score 39; DB 2; Length 525;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPQPQWF 9  
: | | | |  
DB 173 RPNRAWF 181

RESULT 65  
Q9PUA3 PRELIMINARY; PRT; 778 AA.  
ID Q9PUA3;  
AC Q9PUA3;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CIRCADIAN RHYTHMICITY PROTEIN CLOCK.  
OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RT Kim J.S., Drysdale T.A.;  
RT "Sequencing of Xenopus circadian clock gene, xClock.";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF203107; AAF12827.1; -.  
DR HSSP: P36956; 1AM9.  
DR InterPro: IPR001092; HLH\_dim.  
DR InterPro: IPR003015; HLH\_Myc.  
DR InterPro: IPR001067; Nucleinslocator.  
DR InterPro: IPR001610; PAC.  
DR InterPro: IPR000014; PAS.  
DR Pfam: PF00010; HLH; 1.  
DR Pfam: PF00785; PAC; 1.  
DR Pfam: PF00989; PAS; 2.  
DR PRINTS: PR00785; NCTRNLOCATR.  
DR SMART: SM00353; HLH; 1.  
DR SMART: SM00086; PAC; 1.  
DR SMART: SM00091; PAS; 2.  
DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
SQ SEQUENCE 778 AA; 87976 MW; A4D609E88A5F35C4 CRC64;

Query Match 54.9%; Score 39; DB 13; Length 778;  
Best Local Similarity 75.0%; Pred. No. 2.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQQFW 10  
: | | | |  
DB 335 KGQQW 342

RESULT 66  
Q9I906 PRELIMINARY; PRT; 825 AA.  
ID Q9I906;  
AC Q9I906;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CLOCK.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=RETINA.  
RA Zhu H., LaRue S., Whitely A., Steeves T.D.L., Takahashi J.S.,  
RA Green C.B.;  
RT "The Xenopus Clock gene is constitutively expressed in retinal  
photoreceptors.";  
RL Brain Res. Mol. Brain Res. 0:0-0(2000).  
DR EMBL: AF227985; AAF34772.1; -.  
DR InterPro: IPR001092; HLH\_dim.  
DR InterPro: IPR003015; HLH\_Myc.  
DR InterPro: IPR001067; Nucleinslocator.  
DR InterPro: IPR001610; PAC.  
DR InterPro: IPR000014; PAS.  
DR Pfam: PF00785; PAC; 1.  
DR Pfam: PF00989; PAS; 2.  
DR PRINTS: PR00785; NCTRNLOCATR.  
DR SMART: SM00353; HLH; 1.  
DR SMART: SM00086; PAC; 1.  
DR SMART: SM00091; PAS; 2.  
DR PROSITE: PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
SQ SEQUENCE 825 AA; 93269 MW; 0F657C3D9C5F24E0 CRC64;

Query Match 54.9%; Score 39; DB 13; Length 825;  
Best Local Similarity 75.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFNL 10  
| | | | |  
Db 335 KGQOWIWL 342

## RESULT 67

Q9U6M1 PRELIMINARY; PRT; 827 AA.  
AC Q9U6M1;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)  
DE J-BINDING PROTEIN (FRAGMENT).  
GN JBPI.  
OS Leishmania tarentolae (Saurouleishmania tarentolae).  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TAR IVA;  
RX MEDLINE=20031658; PubMed=10562569;  
RA Cross M., Kieft R., Sabatini R., Wilim M., de Kort M.,  
RA van der Marel G.A., van Boom J.H., van Leeuwen F., Borst P.;  
RT "The modified base J is the target for a novel DNA-binding in  
RT kinetoplastid protozoans";  
RL EMBO J. 18:6573-6581(1999).  
DR EMBL; AF182401; AAF01743.1; -.  
FT NON\_TER 827  
SQ SEQUENCE 827 AA; 93403 MW; 5A4B502DE70A4BFE CRC64;

Query Match 54.9%; Score 39; DB 5; Length 827;  
Best Local Similarity 60.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKQOWFWLM 11  
| | : | | | |  
Db 485 PKEEQAFWM 494

## RESULT 68

Q9IAI1 PRELIMINARY; PRT; 852 AA.  
AC Q9IAI1;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE BHLH/PAS TRANSCRIPTION FACTOR CLOCK.  
GN CLOCK.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Noakes M.A., Campbell M., Van Hest B.J.;  
RT "A Chicken Clock Gene Homolog";  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF201076; AAF26365.1; -.  
DR InterPro; IPR001092; HLH\_dlm.  
DR InterPro; IPR003015; HLH\_Myc.  
DR InterPro; IPR001067; Nucleosylator.  
DR InterPro; IPR001610; PAC.  
DR InterPro; IPR000014; PAS.  
DR Pfam; PF00785; PAC; 1.  
DR PRINTS; PR00785; PAC; 2.  
DR

DR SMART; SM00353; HLH; 1.  
DR SMART; SM00086; PAC; 1.  
DR SMART; SM00091; PAS; 2.  
DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
SQ SEQUENCE 852 AA; 96072 MW; 2B4697D7DA72CDOB CRC64;

Query Match 54.9%; Score 39; DB 13; Length 852;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFNL 10  
| | | | |  
Db 344 KGQOWIWL 351

## RESULT 69

Q9W6Q2 PRELIMINARY; PRT; 853 AA.  
AC Q9W6Q2;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE CLOCK PROTEIN.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Chong N.W.S., Klein D.C.;  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF144425; AAD32860.1; -.  
DR HSP; P36956; IAW9.  
DR InterPro; IPR001092; HLH\_dlm.  
DR InterPro; IPR003015; HLH\_Myc.  
DR InterPro; IPR001067; Nucleosylator.  
DR InterPro; IPR001610; PAC.  
DR InterPro; IPR000014; PAS.  
DR Pfam; PF00785; PAC; 1.  
DR Pfam; PF00989; PAC; 2.  
DR PRINTS; PR00785; NCTNSLOCATR.  
DR SMART; SM00353; HLH; 1.  
DR SMART; SM00086; PAC; 1.  
DR SMART; SM00091; PAS; 2.  
DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.  
SQ SEQUENCE 853 AA; 96297 MW; 5349C5C1F7293C97 CRC64;

Query Match 54.9%; Score 39; DB 13; Length 853;  
Best Local Similarity 75.0%; Pred. No. 2.8e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFNL 10  
| | | | |  
Db 344 KGQOWIWL 351

## RESULT 70

Q93244 PRELIMINARY; PRT; 854 AA.  
AC Q93244;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE ANDROGEN RECEPTOR ALPHA.  
GN AR-ALPHA.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;



RL Brain Res. Mol. Brain Res. 0:0-0(1999).

DR EMBL; AF132531; AAD43283.1; -.

DR HSSP; P36956; 1AM9.

DR InterPro; IPR001092; HLH\_dim.

DR InterPro; IPR003015; NucleinsLocator.

DR InterPro; IPR001067; NucleinsLocator.

DR InterPro; IPR001610; PAC.

DR InterPro; IPR000014; PAS.

DR Pfam; PF00785; PAC; 1.

DR Pfam; PF00989; PAC; 2.

DR PRINTS; PR00785; NCTNSLOCATR.

DR SMART; SM00353; HLH; 1.

DR SMART; SM00086; PAC; 1.

DR SMART; SM00091; PAS; 2.

DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.

SQ SEQUENCE 875 AA; 98725 MW; 04DFDEB1D79747A4 CRC64;

Query Match 54.9%; Score 39; DB 13; Length 875;

Best Local Similarity 75.0%; Pred. No. 2.9e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQFWL 10

Db 344 KGOQWIWL 351

RESULT 74

Q9W6J4

ID Q9W6J4

AC Q9W6J4

DT 01-NOV-1999 (TEMBLrel. 12, Created)

DT 01-NOV-1999 (TEMBLrel. 12, Last sequence update)

DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)

DE TRANSCRIPTION FACTOR CLOCK.

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Rasbora; Danio.

OX NCBI\_TaxID=7955;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-99212319; PubMed-10196586;

RA Whitmore D., Foulkes N.S., Strahle U., Sassone-Corsi P.;

RT "zebrafish clock rhythmic expression reveals independent peripheral

circadian oscillators";

RL Nat. Neurosci. 1:701-707(1998).

RN [2]

RP SEQUENCE FROM N.A.

RA Whitmore D., Foulkes N.S., Strahle U., Sassone-Corsi P.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF133306; AAD27749.1; -.

DR InterPro; IPR000014; PAS.

DR InterPro; IPR001067; NucleinsLocator.

DR InterPro; IPR001092; HLH\_dim.

DR InterPro; IPR001610; PAC.

DR InterPro; IPR003015; HLH\_Myc.

DR Pfam; PF00010; HLH; 1.

DR Pfam; PF00785; PAC; 1.

DR Pfam; PF00989; PAC; 2.

DR PRINTS; PR00785; NCTNSLOCATR.

DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN\_1.

DR SMART; SM00353; HLH; 1.

DR SMART; SM00086; PAC; 1.

DR SMART; SM00091; PAS; 2.

SQ SEQUENCE 893 AA; 100945 MW; B74778854457E14A CRC64;

Query Match 54.9%; Score 39; DB 13; Length 893;

Best Local Similarity 75.0%; Pred. No. 3e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQFWL 10

Db 335 KGOQWIWL 342

RESULT 75

Q9VSB0

ID Q9VSB0

AC Q9VSB0

DT 01-MAY-2000 (TEMBLrel. 13, Created)

DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)

DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)

DE CLK PROTEIN.

GN CLK OR CG7391.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=BERKELEY;

RA MEDLINE-20196006; PubMed-10731132;

RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Adair J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,

RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Berson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Burtis K.C., Busay D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harria M.,

RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Lilang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleab J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of Drosophila melanogaster";

RL Science 287:2185-2195(2000).

DR EMBL; AE003557; AAF50516.1; -.

DR FlyBase; FBgn0023076; CLK.

DR InterPro; IPR001092; HLH\_dim.

DR InterPro; IPR003015; HLH\_Myc.

DR InterPro; IPR001067; NucleinsLocator.

DR InterPro; IPR001610; PAC.

DR InterPro; IPR000014; PAS.

DR Pfam; PF00010; HLH; 1.

DR Pfam; PF00785; PAC; 1.

DR Pfam; PF00989; PAC; 2.

DR PRINTS; PR00785; NCTNSLOCATR.

DR SMART; SM00353; HLH; 1.

DR SMART; SM00086; PAC; 1.



DR SMART; SM00091; PAS; 2;  
DR PROSITE; PS00038; HELIX\_LOOP\_HELIX; UNKNOWN 1;  
SQ SEQUENCE 1023 AA; 115691 MW; D4D271038C0D536D CRC64;

Query Match 54.9%; Score 39; DB 5; Length 1023;  
Best Local Similarity 75.0%; Pred. No. 3.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQFWL 10  
:||||  
Db 331 KGOQWIL 338

## RESULT 76

ID Q53785 PRELIMINARY; PRT; 1219 AA.  
AC Q53785;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE TRANSMEMBRANE PROTEIN.  
GN WHB2.  
OS Streptomyces aureofaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1894;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=98234302; PubMed=9565673;  
RA Homerova D., Sprusansky O., Potuckova L., Sevcikova B., Novakova R.,  
RA Rezuchova B., Kormanec J.;  
RT "The gene downstream of Streptomyces aureofaciens whbB encodes a large  
RT protein with proposed transmembrane localization, and is induced by  
RT glucose.";  
RL Biochim. Biophys. Acta 1397:151-155(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Kormanec J., Gabaajova R., Homerova D., Rezuchova B.;  
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.  
DR EMBL; L22864; AAC18892.1; -.  
DR InterPro; IPR001173; Glycos\_transf\_2.  
DR Pfam; PF00535; Glycos\_transf\_2; 1.  
KW Transmembrane.  
SQ SEQUENCE 1219 AA; 128208 MW; 0CEA4C9617ACED32 CRC64;

Query Match 54.9%; Score 39; DB 2; Length 1219;  
Best Local Similarity 55.6%; Pred. No. 4e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQFWL 11  
:||||  
Db 136 EPVQWL 144

## RESULT 77

ID Q9FVG4 PRELIMINARY; PRT; 673 AA.  
AC Q9FVG4;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE TRANSPOSASE DOPA.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=W22; TRANSPOSON=DOPPIA;  
RA Bercury S.D., Walker E.L.;

RT "Molecular analysis of the structure and function of the Doppia  
RT transposable element of maize.";  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF187823; AAG17044.1; -.  
SQ SEQUENCE 673 AA; 75138 MW; 46241159B444D59D CRC64;

Query Match 54.2%; Score 38.5; DB 10; Length 673;  
Best Local Similarity 60.0%; Pred. No. 2.7e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPK-PQOWFW 9  
||:|:|  
Db 623 RPRLPRPWF 632

## RESULT 78

Q9HZ24  
ID Q9HZ24 PRELIMINARY; PRT; 105 AA.  
AC Q9HZ24;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL PROTEIN PA3216.  
GN PA3216.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Miziouchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
DR EMBL; AE004745; AAG06604.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 105 AA; 12396 MW; D98630BC71A85177 CRC64;

Query Match 53.5%; Score 38; DB 2; Length 105;  
Best Local Similarity 71.4%; Pred. No. 53;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 QQFWLM 11  
|||  
Db 50 QQWLWL 56

## RESULT 79

ID Q62024 PRELIMINARY; PRT; 122 AA.  
AC Q62024; P97339;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE PP52 (FRAGMENT).  
GN LSP1 OR PP52.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C;  
RX MEDLINE=96435912; PubMed=8838798;

RA Thompson A.A., Omori S.A., Gilly M.J., May W., Gordon M.S., Wood W.J.,  
 RA Miyoshi E., Malone C.S., Gimble J., Kincade P.W., Denny C.T., Wall R.;  
 RT "Alternatively spliced exons encode the tissue-specific 5' termini of  
 RT leukocyte pp52 and stromal cell S37 mRNA isoforms.";  
 RL Genomics 32:352-357(1996).  
 DR EMBL; U30942; AAB37543.1;  
 DR EMBL; U30940; AAB37543.1; JOINED.  
 DR EMBL; U30941; AAB37543.1; JOINED.  
 DR MGD; MGI:96832; Lspl.  
 DR InterPro; IPR002211; Lymphspecific.  
 DR PRINTS; PR01083; LYMPHSPECIFIC.  
 FT NON\_TER 122 122  
 SQ SEQUENCE 122 AA; 14237 MW; AB6AA7AF4FE7D3B1 CRC64;

Query Match 53.5%; Score 38; DB 11; Length 122;  
 Best Local Similarity 55.6%; Pred. No. 61;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
 :|:|:|  
 Db 90 KPEPRQWF 98

RESULT 80

Q62023 PRELIMINARY; PRT; 124 AA.  
 AC Q62023;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE PP52 PROTEIN (FRAGMENT).  
 GN LSP1 OR PP52.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C;  
 RX MEDLINE=96435912; PubMed=8838798;  
 RA Thompson A.A., Omori S.A., Gilly M.J., May W., Gordon M.S., Wood W.J.,  
 RA Miyoshi E., Malone C.S., Gimble J., Kincade P.W., Denny C.T., Wall R.;  
 RT "Alternatively spliced exons encode the tissue-specific 5' termini of  
 RT leukocyte pp52 and stromal cell S37 mRNA isoforms.";  
 RL Genomics 32:352-357(1996).  
 DR EMBL; U30942; AAB37542.1;  
 DR EMBL; U30939; AAB37542.1; JOINED.  
 DR EMBL; U30941; AAB37542.1; JOINED.  
 DR MGD; MGI:96832; Lspl.  
 DR InterPro; IPR002211; Lymphspecific.  
 DR PRINTS; PR01083; LYMPHSPECIFIC.  
 FT NON\_TER 124 124  
 SQ SEQUENCE 124 AA; 14388 MW; D98E59E2EF81954F CRC64;

Query Match 53.5%; Score 38; DB 11; Length 124;  
 Best Local Similarity 55.6%; Pred. No. 62;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
 :|:|:|  
 Db 92 KPEPRQWF 100

RESULT 81

Q9T6A9 PRELIMINARY; PRT; 175 AA.  
 AC Q9T6A9;  
 DT 01-MAY-2000 (TReMBLrel. 13, Created)  
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1) (FRAGMENT).  
 GN

GN COI.  
 OS Culicoides tuttifrutti.  
 OC Mitochondrion.  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;  
 OC Chironomidae; Ceratopogonidae; Culicoides.  
 ON NCBI\_TaxID=88402;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TUT 1;  
 RA Linton Y.-M., Meiswinkel R., Venter G.J., Mellor P.S.,  
 RA "Phylogenetic relationship between five members of the Culicoides  
 RT mydologene complex in South Africa based on mtDNA cytochrome  
 RT oxidase I sequence data.";  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY  
 CC CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-  
 CC 3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. COI IS THE  
 CC CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN  
 CC CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2  
 CC AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3  
 CC AND COPPER B (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O + 4  
 CC FERRICYTOCHROME C.  
 CC -!- COFACTOR: TWO HEME (A AND A3) GROUPS AND COPPER B (BY SIMILARITY).  
 CC -!- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN  
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.  
 DR EMBL; AF069242; AAD32471.1;  
 DR InterPro; IPR002106; AA\_TRNA\_ligase\_II.  
 DR InterPro; IPR000883; COXI.  
 DR InterPro; IPR001360; Glyco\_hydro\_1.  
 DR Pfam; PF00115; COXI; 1.  
 DR PRINTS; PR01165; CYCOXIDASEI.  
 DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; UNKNOWN\_1.  
 DR PROSITE; PS00572; GLYCOSYL\_HYDROL\_F1\_1; UNKNOWN\_1.  
 KW Copper; Heme; Inner membrane; Mitochondrion; Oxidoreductase;  
 KW Respiratory chain; Transmembrane.  
 FT NON\_TER 175 175  
 FT NON\_TER 175 175  
 SQ SEQUENCE 175 AA; 18767 MW; 2D5A93BBDF049DD3 CRC64;

Query Match 53.5%; Score 38; DB 8; Length 175;  
 Best Local Similarity 100.0%; Pred. No. 87;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QQWF 9  
 :|:|:|  
 Db 157 QQWF 161

RESULT 82

Q9D1Z1 PRELIMINARY; PRT; 188 AA.  
 AC Q9D1Z1;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
 DE ADULT RETINA CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
 DE CLONE:A930019G03, FULL INSERT SEQUENCE.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=RETINA;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Harsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaudo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Nombarts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK020876; BAB32236.1; -;  
 SQ SEQUENCE 188 AA; 19650 MW; FFB762B7A4F83F1A CRC64;

Query Match 53.5%; Score 38; DB 11; Length 188;  
 Best Local Similarity 50.0%; Pred. No. 93;  
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PKPQOWFLM 11  
 | | | | |  
 DB 156 PGSDSWFWLL 165

RESULT 83  
 O07878  
 ID O07878 PRELIMINARY; PRT; 220 AA.  
 AC O07878;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE FDX1 GENE AND 4 ORF'S.  
 OS Spingomonas sp.  
 CC Bacteria; Proteobacteria; alpha subdivision; Sphingomonas group;  
 CC Sphingomonas.  
 OC NCBI\_TaxID=28214;  
 OX [1]  
 RN RP SEQUENCE FROM N.A.  
 RC STRAIN=RW1;  
 RX MEDLINE=97433265; PubMed=9288905;  
 RA Armenaud J., Timmis K.N.;  
 RT "Molecular characterization of Fdx1, a putidaredoxin-type [2Fe-2S]  
 ferredoxin able to transfer electrons to the dioxin dioxygenase of  
 RT Sphingomonas sp. RW1.";  
 RL Eur. J. Biochem. 247:833-842(1997).  
 DR EMBL; Y13118; CAA73584.1; -;  
 DR InterPro; IPR000521; GST.  
 DR Pfam; PF00043; GST; 2.  
 SQ SEQUENCE 220 AA; 24568 MW; 47C3C2BF66519948 CRC64;

Query Match 53.5%; Score 38; DB 2; Length 220;  
 Best Local Similarity 44.4%; Pred. No. 11e+02;  
 Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RKPKQOWFW 9  
 | | | | |  
 DB 210 QPDSEFWF 218

RESULT 84  
 Q9ANCO  
 ID Q9ANCO PRELIMINARY; PRT; 253 AA.  
 AC Q9ANCO;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE ID342.  
 GS ID342.  
 OS Bradyrhizobium japonicum.  
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC Bradyrhizobium group; Bradyrhizobium.  
 OX NCBI\_TaxID=375;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=110SPC4;  
 RX MEDLINE=21101824; PubMed=11157954;  
 RA Gottfert M., Rothlisberger S., Kundig C., Beck C., Marty R.,  
 RA Hennecke H.;  
 RT "Potential symbiosis-specific genes uncovered by sequencing a 410-kb  
 RT DNA region of the Bradyrhizobium japonicum chromosome.";  
 RL J. Bacteriol. 183:1405-1412(2001).  
 DR EMBL; AF322012; AAG60852.1; -;  
 SQ SEQUENCE 253 AA; 27851 MW; 88E7222D98506C13 CRC64;

Query Match 53.5%; Score 38; DB 2; Length 253;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RKPKQOW 7  
 | | | | |  
 DB 201 RPLPQOW 207

RESULT 85  
 Q9GUP1  
 ID Q9GUP1 PRELIMINARY; PRT; 260 AA.  
 AC Q9GUP1;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL PROTEIN Y67D8C.H.  
 GN Y67D8C.H.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 OC NCBI\_TaxID=6239;  
 OX [1]  
 RN RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BRISTOL N2;  
 RA Waterston R.;  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC025724; AAG23376.1; -;  
 SQ SEQUENCE 260 AA; 30643 MW; 78CADE13E2AAA59A CRC64;

Query Match 53.5%; Score 38; DB 5; Length 260;  
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQOWFLM 11  
 | | | | |  
 DB 175 PSQWLWL 182

RESULT 86  
 P79391  
 ID P79391 PRELIMINARY; PRT; 270 AA.  
 AC P79391;  
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)

DT 01-JUN-2001 (TREMELREL. 17, Last annotation update)  
 DE LECTIN-LIKE OXIDIZED LDL RECEPTOR.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=97205278; PubMed=9052782;  
 RA Sawamura T., Kume N., Aoyama T., Moriaki H., Hoshikawa H., Aiba Y.,  
 RA Tanaka T., Miwa S., Katsura Y., Kita T., Masaki T.,  
 RA "An endothelial receptor for oxidized low-density lipoprotein."  
 RL Nature 386:73-77(1997).  
 DR EMBL; D89049; BAA19005.1; .  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR PROSITE; PS50041; C\_TYPE\_LLECTIN\_2; 1.  
 DR SMART; SM00034; CLECT; 1.  
 KW Lectin.  
 SQ SEQUENCE 270 AA; 30892 MW; 6055B6881AD7053D CRC64;  
  
 Query Match 53.5%; Score 38; DB 6; Length 270;  
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
 QY 2 PKPQWFW 9  
 Db 139 PCPDWLW 146  
  
 RESULT 87  
 P78380  
 ID P78380 PRELIMINARY; PRT; 273 AA.  
 AC P78380;  
 DT 01-MAY-1997 (TREMELREL. 03, Created)  
 DT 01-MAY-1997 (TREMELREL. 03, Last sequence update)  
 DT 01-JUN-2001 (TREMELREL. 17, Last annotation update)  
 DE LECTIN-LIKE OXIDIZED LDL RECEPTOR.  
 GN LOX-1 OR OLR1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=LUNG;  
 RC MEDLINE=97205278; PubMed=9052782;  
 RX Sawamura T., Kume N., Aoyama T., Moriaki H., Hoshikawa H., Aiba Y.,  
 RA Tanaka T., Miwa S., Katsura Y., Kita T., Masaki T.,  
 RA "An endothelial receptor for oxidized low-density lipoprotein."  
 RL Nature 386:73-77(1997).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=BLOOD;  
 RC Millar D.S.;  
 RA Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RP Li X., Bouzyk M.M., Wang X.K.;  
 RA "Human oxidized low density lipoprotein receptor: characterization of  
 RT the full length cDNA sequence and assignment to human chromosome  
 RT 12p13.1-12.3."  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=99047525; PubMed=9828121;  
 RX Yamanaka S., Zhang X.Y., Miura K., Kim S., Iwao H.;  
 RA "The human gene encoding the lectin-type oxidized LDL receptor (OLR1)  
 RT is a novel member of the natural killer gene complex with a unique  
 RT expression profile."  
 RN Genomics 54:191-199(1998).

DR EMBL; AB010710; BAA24580.1; .  
 DR EMBL; AJ131757; CAB38175.1; .  
 DR EMBL; AF035776; AAC82329.1; .  
 DR EMBL; AF079167; AAC97927.1; .  
 DR EMBL; AF079164; AAC97927.1; JOINED.  
 DR EMBL; AF079165; AAC97927.1; JOINED.  
 DR EMBL; AF079166; AAC97927.1; JOINED.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR PROSITE; PS50041; C\_TYPE\_LLECTIN\_2; 1.  
 DR SMART; SM00034; CLECT; 1.  
 KW Lectin; Receptor; Lipoprotein.  
 SQ SEQUENCE 273 AA; 30959 MW; 852DE6595DC3D361 CRC64;  
  
 Query Match 53.5%; Score 38; DB 4; Length 273;  
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
 QY 2 PKPQWFW 9  
 Db 143 PCPDWLW 150  
  
 RESULT 88  
 Q9TTK7  
 ID Q9TTK7 PRELIMINARY; PRT; 274 AA.  
 AC Q9TTK7;  
 DT 01-MAY-2000 (TREMELREL. 13, Created)  
 DT 01-MAY-2000 (TREMELREL. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMELREL. 17, Last annotation update)  
 DE LECTIN-LIKE OXIDIZED LDL RECEPTOR-1.  
 GN PLOX-1.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP Sawamura T.;  
 RA "Porcine lectin-like oxidized LDL receptor-1(LOX-1).";  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB018668; BAA89894.1; .  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR PROSITE; PS50041; C\_TYPE\_LLECTIN\_2; 1.  
 DR SMART; SM00034; CLECT; 1.  
 KW Receptor; Lectin.  
 SQ SEQUENCE 274 AA; 31142 MW; D141776C79FB42E0 CRC64;  
  
 Query Match 53.5%; Score 38; DB 6; Length 274;  
 Best Local Similarity 62.5%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
 QY 2 PKPQWFW 9  
 Db 143 PCPDWLW 150  
  
 RESULT 89  
 Q99L65  
 ID Q99L65 PRELIMINARY; PRT; 322 AA.  
 AC Q99L65;  
 DT 01-JUN-2001 (TREMELREL. 17, Created)  
 DT 01-JUN-2001 (TREMELREL. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMELREL. 17, Last annotation update)  
 DE SIMILAR TO LYMPHOCYTE SPECIFIC 1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE-MAMMARY TUMOR;  
RA Strausberg R.;  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC003796; AA003796.1; -;  
SQ SEQUENCE 322 AA; 35763 MW; 59D77C19CDE27651 CRC64;

Query Match 53.5%; Score 38; DB 11; Length 322;  
Best Local Similarity 55.6%; Pred. No. 1.6e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 9  
:|:|:|  
Db 90 KPEPRQFW 98

RESULT 90  
Q62022 PRELIMINARY; PRT; 330 AA.  
AC Q62022;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE P50B.  
GN LSP1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=ICR; TISSUE=THYMUS;  
RA MEDLINE=95293928; PubMed=7775393;  
RA Matsumoto N., Kojima S., Osawa T., Toyoshima S.;  
RT "Protein kinase C phosphorylates p50 LSP1 and induces translocation of  
p50 LSP1 in T lymphocytes.";  
RL J. Biochem. 117:222-229(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ICR; TISSUE=THYMUS;  
RA MEDLINE=96015175; PubMed=8537319;  
RA Matsumoto N., Kita K., Kojima S., Yamamoto K., Irimura T., Miyagi M.,  
RA Tsunawasa S., Toyoshima S.;  
RT "Lymphocyte isoforms of mouse p50 LSP1, which are phosphorylated in  
mitogen-activated T cells, are formed through alternative splicing and  
phosphorylation.";  
RL J. Biochem. 118:237-243(1995).  
DR EMBL; D49691; BAA08541.1; -;  
DR MGD; MGI:96832; Lsp1.  
DR InterPro; IPR002211; Lymphspecific.  
DR PRINTS; PR01083; LYMPHSFCIFC.  
SQ SEQUENCE 330 AA; 36728 MW; 3CC27400F02859FD CRC64;

Query Match 53.5%; Score 38; DB 11; Length 330;  
Best Local Similarity 55.6%; Pred. No. 1.6e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 9  
:|:|:|  
Db 92 KPEPRQFW 100

RESULT 91  
Q98952 PRELIMINARY; PRT; 356 AA.  
AC Q98952;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE ALPHA-(1.3)-FUCOSYLTRANSFERASE (EC 2.4.1.-)  
DE (GALACTOSIDE 3-L-FUCOSYLTRANSFERASE) (FUCT-IV)

DE (CTF1).  
GN CFT1.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=97115837; PubMed=89551139;  
RA Lee K.P., Carlson L.M., Woodcock J.B., Ramachandra N., Schultz T.L.,  
RA Davis T.A., Lowe J.B., Thompson C.B., Larsen R.D.;  
RL J. Biol. Chem. 271:32960-32967(1996).  
CC -!- FUNCTION: MAY CATALYSE ALPHA-1,3 GLYCOSIDIC LINKAGES INVOLVED IN  
THE EXPRESSION OF LEWIS X/SSEA-1 AND VIM-2 ANTIGENS.  
CC -!- PATHWAY: GLYCOSYLATION.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
FORM IN TRANS CISTERNAE OF GOLGI.  
CC -!- TISSUE SPECIFICITY: IN THE FOLLOWING EMBRYONIC TISSUES: BRAIN,  
EYE, GIZZARD, THYMUS, BURSA AND SPLEEN.  
DR EMBL; U73678; AAC60060.1; -;  
DR InterPro; IPR001503; Glyco\_transf\_10.  
DR Pfam; PF00852; Glyco\_transf\_10; 1.  
KW Transferase; Glycosyltransferase; Transmembrane; Glycoprotein;  
KW Signal-anchor; Golgi stack.  
FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 23 51 SIGNAL-ANCHOR (POTENTIAL).  
FT DOMAIN 52 356 (TYPE-II MEMBRANE PROTEIN).  
FT CARBOHYD 80 80 LUMENAL, CATALYTIC (POTENTIAL).  
FT CARBOHYD 149 149 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 149 149 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 356 AA; 41494 MW; 13141627FE8AD089 CRC64;

Query Match 53.5%; Score 38; DB 13; Length 356;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 10  
:|:|:|:|  
Db 121 RPPRQRWVWM 130

RESULT 92  
Q9EQ09 PRELIMINARY; PRT; 363 AA.  
ID Q9EQ09  
AC Q9EQ09;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE OXIDIZED LDL RECEPTOR.  
GN LOX-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Park S.-H., Ahn H.-J., Cho J.-J.;  
RT "Mouse LOX-1 is expressed in mast cells after IgE cross-linking.";  
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF303744; AAG44998.1; -;  
DR InterPro; IPR001304; lectin\_c.  
DR Pfam; PF00059; lectin\_c; 1.  
DR SMART; SM00034; CLECT; 1.  
DR PROSITE; PS50041; C\_TYPE\_LLECTIN\_2; 1.  
KW Receptor.  
SQ SEQUENCE 363 AA; 41613 MW; E44703D6408F15F8 CRC64;

Query Match 53.5%; Score 38; DB 11; Length 363;  
Best Local Similarity 62.5%; Pred. No. 1.8e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
QY 2 PKPQQWF 9
   |||
DB 234 PCQDWLW 241

RESULT 93
ID 070156 PRELIMINARY; PRT; 364 AA.
AC O70156
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ENDOTHelial RECEPTOR FOR OXIDIZED LOW-DENSITY LIPOPROTEIN.
GN LOX-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SHR-SP; TISSUE=KIDNEY;
RX MEDLINE=98161826; PubMed=9494115;
RA Nagase M., Hirose S., Fujita T.;
RT "Unique repetitive sequence and unexpected regulation of expression of
RT rat endothelial receptor for oxidized low-density lipoprotein (LOX-
RT 1).";
RL Biochem. J. 330:1417-1422(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RX MEDLINE=99057940; PubMed=9837956;
RA Nagase M., Abe J., Takahashi K., Ando J., Hirose S., Fujita T.;
RT "Genomic organization and regulation of expression of the lectin-like
RT oxidized low-density lipoprotein receptor (LOX-1) gene.";
RL J. Biol. Chem. 273:33702-33707(1998).
DR ENBL; AB005900; BAA25785.1; -.
DR ENBL; AB018104; BAA35123.1; -.
DR ENBL; AB018097; BAA35123.1; JOINED.
DR ENBL; AB018098; BAA35123.1; JOINED.
DR ENBL; AB018099; BAA35123.1; JOINED.
DR ENBL; AB018100; BAA35123.1; JOINED.
DR ENBL; AB018101; BAA35123.1; JOINED.
DR ENBL; AB018102; BAA35123.1; JOINED.
DR ENBL; AB018103; BAA35123.1; JOINED.
DR InterPro: IPR001304; Lectin_c.
DR Pfam: PF00059; lectin_c; 1.
DR PROSITE: PSS0041; C-TYPE-LECTIN_2; 1.
DR SMART: SM00034; CLECT; 1.
KW Lipoprotein; Receptor; Lectin.
SQ SEQUENCE 364 AA; 41890 MW; 0AD2839C07206E09 CRC64;

Query Match 53.5%; Score 38; DB 11; Length 364;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQWF 9
   |||
DB 234 PCQDWLW 241

RESULT 94
ID 09A8W6 PRELIMINARY; PRT; 368 AA.
AC 09A8W6
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN CC1232.
GN CC1232.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
```

```
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Padake N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vanathavan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005799; AAK23214.1; -.
DR TIGR; CC1232; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 368 AA; 41087 MW; 6985682449AD7A29 CRC64;

Query Match 53.5%; Score 38; DB 2; Length 368;
Best Local Similarity 66.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQQWF 10
   |||
DB 194 PVPPQFWL 202

RESULT 95
ID 09PDR6 PRELIMINARY; PRT; 373 AA.
AC 09PDR6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ROD SHAPE-DETERMINING PROTEIN.
GN Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hoheisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peixoto B.R., Pereira C.A.G., Pereira H.A. Jr., Pesquero J.B.,
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
```

```

RL Nature 406:151-159(2000).
DR EMBL: AE003964; AAF84122.1; -.
DR InterPro: IPR001182; FTSW_RodA_SPOVE.
DR Pfam: PF01098; FTSW_RodA_SPOVE; 1.
DR PROSITE: PS00428; FTSW_RodA_SPOVE; 1.
KW Complete proteome.
SQ SEQUENCE 373 AA; 41436 MW; E9480421BF70EAC5 CRC64;

Query Match 53.5%; Score 38; DB 2; Length 373;
Best Local Similarity 62.5%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PQQWFLL 11
Db 201 PFSNFWLL 208

RESULT 96
Q9FHE8 PRELIMINARY; PRT; 392 AA.
AC Q9FHE8;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE SIMILARITY TO DISEASE RESISTANCE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
ON NCBI_TaxID=3702;
RX Medline=20181125; Pubmed=10718197;
RC STRAIN=COLUMBIA;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
RT features of the regions of 3,076,755 bp covered by sixty P1 and TAC
RT clones."
RL DNA Res. 7:31-63(2000).
DR EMBL: AB019224; BAB09491.1; -.
DR InterPro: IPR000157; TIR.
DR Pfam: PF01582; TIR; 1.
DR SMART: SM00255; TIR; 1.
SQ SEQUENCE 392 AA; 44141 MW; 8E8730F454878CA2 CRC64;

Query Match 53.5%; Score 38; DB 10; Length 392;
Best Local Similarity 71.4%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KPQQWF 9
Db 251 KPQKWTW 257

RESULT 97
Q9JIK2 PRELIMINARY; PRT; 401 AA.
AC Q9JIK2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE ALPHA-(1,3)-FUCOSYLTRANSFERASE (EC 2.4.1.-) (GALACTOSIDE 3-L-
DE FUCOSYLTRANSFERASE) (FUCOSYLTRANSFERASE 4) (FUCT-IV).
GN FUT4.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
ON NCBI_TaxID=10029;
RX Medline=10029;
SQ SEQUENCE 401 AA; 45309 MW; 3130298017560D13 CRC64;

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=OVARY.
RX MEDLINE=20166953; Pubmed=10700388;
RA Patnaik S.K., Zhang A., Shi S., Stanley P.;
RT "Alpha(1,3)fucosyltransferases expressed by the gain-of-function
RT Chinese hamster ovary glycosylation mutants LEC12, LEC29, and LEC30."
RL Arch. Biochem. Biophys. 375:322-332(2000).
CC -!- FUNCTION: MAY CATALYZE ALPHA-1,3 GLYCOSIDIC LINKAGES INVOLVED IN
CC THE EXPRESSION OF LEWIS X AND LEWIS Y.
CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + 1,4-BETA-D-GALACTOSYL-N-
CC ACETYL-D-GLUCOSAMINYL-R = GDP + 1,4-BETA-D-GALACTOSYL-(ALPHA-
CC 1,3-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINYL-R.
CC -!- PATHWAY: GLYCOSYLATION.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
DR EMBL: AF221505; AAF82352.1; -.
DR InterPro: IPR001503; Glyco_transf_10.
DR Pfam: PF00852; Glyco_transf_10; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Glycoprotein;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 24 42 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 43 401 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 401 AA; 45309 MW; 3130298017560D13 CRC64;

Query Match 53.5%; Score 38; DB 11; Length 401;
Best Local Similarity 50.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKQQWF 10
Db 157 RPPQQRWVM 166

RESULT 98
Q26064 PRELIMINARY; PRT; 412 AA.
ID Q26064;
AC Q26064;
DT 01-JAN-1998 (Tremblrel. 05, Created)
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE UBIQUINOL CYTOCHROME C OXIDOREDUCTASE, CYTOCHROME B SUBUNIT (FBCH).
GN CYTB OR HP1539.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; Pubmed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).

```

```
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.  
DR EMBL: AE000652; AAD08579.1; -.  
DR TIGR: HP1539; -.  
DR InterPro: IPR000179; Cyt_b_b6.  
DR Pfam: PF00032; cytochrome_b_c1; 1.  
DR Pfam: PF00033; cytochrome_b_n; 1.  
DR PROSITE: PS00193; CYTOCHROME_B_QO; UNKNOWN.1.  
DR Complete proteome; Electron transport; Heme; Hypothetical protein;  
KW Respiratory chain; Transmembrane.  
SQ SEQUENCE 412 AA; 47510 MW; 954646D95A1F925A CRC64;  
  
Query Match 53.5%; Score 38; DB 2; Length 412;  
Best Local Similarity 54.5%; Pred. No. 2e+02;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 RPXQQQFWLM 11  
|||  
DB 349 RPAPMVWFLV 359  
|||  
  
RESULT 99  
Q9CAP2 PRELIMINARY; PRT; 421 AA.  
AC Q9CAP2;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE HYPOTHETICAL 47.7 KDA PROTEIN.  
GN T5M16.25.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV, COLUMBIA;  
RX MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,  
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
RA Sakano H., Salter J.S., Schwartz J.R., Shinn P., Southwick A.M.,  
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana.";  
RL Nature 408:816-820(2000).  
DR EMBL: AC010704; AAG51667.1; -.  
DR InterPro: IPR003409; MORN.  
DR Pfam: PF02493; MORN; 7.  
KW Hypothetical protein.  
SQ SEQUENCE 421 AA; 47731 MW; 08361CB916235663 CRC64;  
  
Query Match 53.5%; Score 38; DB 10; Length 421;  
Best Local Similarity 62.5%; Pred. No. 2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPXQQQFW 8  
:|:|:|  
  
CC -1- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.  
DR EMBL: AE000652; AAD08579.1; -.  
DR TIGR: HP1539; -.  
DR InterPro: IPR000179; Cyt_b_b6.  
DR Pfam: PF00032; cytochrome_b_c1; 1.  
DR Pfam: PF00033; cytochrome_b_n; 1.  
DR PROSITE: PS00193; CYTOCHROME_B_QO; UNKNOWN.1.  
DR Complete proteome; Electron transport; Heme; Hypothetical protein;  
KW Respiratory chain; Transmembrane.  
SQ SEQUENCE 412 AA; 47510 MW; 954646D95A1F925A CRC64;  
  
Query Match 53.5%; Score 38; DB 2; Length 412;  
Best Local Similarity 54.5%; Pred. No. 2e+02;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 RPXQQQFWLM 11  
|||  
DB 349 RPAPMVWFLV 359  
|||  
  
RESULT 99  
Q9CAP2 PRELIMINARY; PRT; 421 AA.  
AC Q9CAP2;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE HYPOTHETICAL 47.7 KDA PROTEIN.  
GN T5M16.25.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV, COLUMBIA;  
RX MEDLINE=21016719; PubMed=11130712;  
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,  
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,  
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,  
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,  
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,  
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,  
RA Lin X., Liu S.X., Liu Z.A., Lueros J.S., Maiti R., Marziani A.,  
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,  
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,  
RA Sakano H., Salter J.S., Schwartz J.R., Shinn P., Southwick A.M.,  
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
thaliana.";  
RL Nature 408:816-820(2000).  
DR EMBL: AC010704; AAG51667.1; -.  
DR InterPro: IPR003409; MORN.  
DR Pfam: PF02493; MORN; 7.  
KW Hypothetical protein.  
SQ SEQUENCE 421 AA; 47731 MW; 08361CB916235663 CRC64;  
  
Query Match 53.5%; Score 38; DB 10; Length 421;  
Best Local Similarity 62.5%; Pred. No. 2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPXQQQFW 8  
:|:|:|
```

```
Db 145 KXPVQWY 152  
  
RESULT 100  
Q99N88 PRELIMINARY; PRT; 433 AA.  
ID Q99N88;  
AC Q99N88;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE ALPHA1,3-FUCOSYLTRANSFERASE.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI_TaxID=101116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sanai Y., Shimoda Y., Tajima Y., Osanai T., Katsume A., Kohara M.,  
RA Kudo T., Narimatsu H., Osumi N.;  
RT "Expression of Lewis x epitope in embryonic forebrain.";  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB049938; BAB40992.1; -.  
KW Transferase; Glycosyltransferase.  
SQ SEQUENCE 433 AA; 48918 MW; 1E1AFFB70EFA1402 CRC64;  
  
Query Match 53.5%; Score 38; DB 11; Length 433;  
Best Local Similarity 50.0%; Pred. No. 2.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPXQQQFWL 10  
|||  
DB 189 RPPQQRWVM 198  
|||  
  
RESULT 101  
O63524 PRELIMINARY; PRT; 455 AA.  
ID O63524;  
AC O63524;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)  
DE HYPOTHETICAL 51.8 KDA PROTEIN.  
GN F4D11.30 OR AT4G32770.  
OS Arabidopsis thaliana (Mouse-ear cross).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bevan M., Benes V., Rechmann S., Borkova D., Ansoorge W., Hoheisel J.,  
RA Mewes H.W., Mayer K.F.X., Schueller C.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Benes V., Rechmann S., Borkova D., Ansoorge W., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL022537; CAA18584.1; -.  
DR EMBL: AL161582; CAB79994.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 455 AA; 51838 MW; 4A05CF5F0BFA994D CRC64;  
  
Query Match 53.5%; Score 38; DB 10; Length 455;  
Best Local Similarity 57.1%; Pred. No. 2.2e+02;  
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
```



Qy 4 PQOWFWL 10  
I:||||  
Db 242 PRKWFV 248

## RESULT 102

Q9WID1 PRELIMINARY: PRT: 458 AA.  
AC Q9WID1;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE CG11388 PROTEIN.  
GN CG11388.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=BERKELEY;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis J.M., Cawley S., Dahlke C., Cadieu E., Center A., Chandra I.,  
RA de Pablo K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Dodson K., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fogle K., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeguam C.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,  
RA Mount S.M., Nixon K.A., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Wang X.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Zheng X.H., Zhong F.N., Zhou W., Zhang G., Zhao Q., Zheng L.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Zheng X.H., Zhong F.N., Zhou W., Zhang G., Zhao Q., Zheng L.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003462; AAF47138.1; -.  
DR FlyBase; FBgn0034959; CG11388.  
SQ SEQUENCE 458 AA: 53563 MW; 673A3F96F7C4BA9A CRC64;

Query Match 53.5%; Score 38; DB 5; Length 458;  
Best Local Similarity 45.5%; Pred. No. 2.2e+02;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RPKPQOWFWL 11  
I:||||  
Db 50 ROKPKSAWTL 60

## RESULT 103

Q9VUK7 PRELIMINARY: PRT: 635 AA.  
AC Q9VUK7;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
DE CG6945 PROTEIN.  
GN CG6945.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=BERKELEY;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis J.M., Cawley S., Dahlke C., Cadieu E., Center A., Chandra I.,  
RA de Pablo K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Dodson K., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fogle K., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeguam C.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nixon K.A., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Wang X.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Zheng X.H., Zhong F.N., Zhou W., Zhang G., Zhao Q., Zheng L.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Zheng X.H., Zhong F.N., Zhou W., Zhang G., Zhao Q., Zheng L.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003532; AAF49669.1; -.  
DR FlyBase; FBgn0036476; CG6945.  
SQ SEQUENCE 635 AA: 70588 MW; 7F02A40BF519E5AC CRC64;

Query Match 53.5%; Score 38; DB 5; Length 635;  
Best Local Similarity 71.4%; Pred. No. 3e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQOW 7  
I:||||  
Db 400 RPKPKQW 406

## RESULT 104

Q9UAC1 PRELIMINARY; PRT; 684 AA.  
AC Q9UAC1;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE PUTATIVE PTERIDINE TRANSPORTER FT2.  
OS Leishmania donovani.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5661;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IS2D;  
RA Moore J.B., Beverley S.M.;  
RT "Identification of a protein mediating folate/methotrexate uptake by  
RT genetic rescue of a Leishmania transport mutant.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF084470; AAD52047.1; -.  
SQ SEQUENCE 684 AA; 75387 MW; 9F773E0DC562A037 CRC64;  
  
Query Match 53.5%; Score 38; DB 5; Length 684;  
Best Local Similarity 43.8%; Pred. No. 3.3e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 6; Gaps 1;  
  
QY 2 PKPQQ-----WFWLM 11  
||| |||:  
Db 193 PKPGPSMVSWIFWIM 208  
  
RESULT 105  
Q9UAB7 PRELIMINARY; PRT; 691 AA.  
AC Q9UAB7;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE PUTATIVE PTERIDINE TRANSPORTER FT6.  
OS Leishmania donovani.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5661;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IS2D;  
RA Moore J.B., Beverley S.M.;  
RT "Identification of a protein mediating folate/methotrexate uptake by  
RT genetic rescue of a Leishmania transport mutant.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF084474; AAD52051.1; -.  
SQ SEQUENCE 691 AA; 76005 MW; 63D1D7671D8B45E6 CRC64;  
  
Query Match 53.5%; Score 38; DB 5; Length 691;  
Best Local Similarity 43.8%; Pred. No. 3.3e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 6; Gaps 1;  
  
QY 2 PKPQQ-----WFWLM 11  
||| |||:  
Db 193 PKPGPSMVSWIFWIM 208  
  
RESULT 106  
Q9LUQ4 PRELIMINARY; PRT; 697 AA.  
AC Q9LUQ4;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE GENOMIC DNA, CHROMOSOME 3, PL CLONE: MK13.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RC MEDLINE=20277480; PubMed=10819329;  
RX Nakamura Y.;  
RA "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
RT features of the regions of 4,504,864 bp covered by sixty pl and TAC  
RT clones.";  
RL DNA Res. 7:131-135(2000).  
DR EMBL; AB022218; BAB02367.1; -.  
DR InterPro; IPR001211; PLP\_A2.  
DR PROSITE; PS00118; PA2\_HIS; UNKNOWN\_1.  
DR SEQUENCE 697 AA; 78550 MW; 421B2F2E6CD39AA7 CRC64;  
  
Query Match 53.5%; Score 38; DB 10; Length 697;  
Best Local Similarity 62.5%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 4 PQQWFWLM 11  
||| |||:  
Db 525 PDHFWFRM 532  
  
RESULT 107  
Q9UAC0 PRELIMINARY; PRT; 698 AA.  
AC Q9UAC0;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE PUTATIVE PTERIDINE TRANSPORTER FT3.  
OS Leishmania donovani.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5661;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IS2D;  
RA Moore J.B., Beverley S.M.;  
RT "Identification of a protein mediating folate/methotrexate uptake by  
RT genetic rescue of a Leishmania transport mutant.";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF084471; AAD52048.1; -.  
SQ SEQUENCE 698 AA; 75779 MW; B817D80B76A1AFDB CRC64;  
  
Query Match 53.5%; Score 38; DB 5; Length 698;  
Best Local Similarity 43.8%; Pred. No. 3.3e+02;  
Matches 7; Conservative 1; Mismatches 2; Indels 6; Gaps 1;  
  
QY 2 PKPQQ-----WFWLM 11  
||| |||:  
Db 194 PKPGPSMVSWIFWIM 209  
  
RESULT 108  
Q9UAB9 PRELIMINARY; PRT; 700 AA.  
AC Q9UAB9;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
DE PUTATIVE PTERIDINE TRANSPORTER FT4.  
OS Leishmania donovani.  
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5661;  
RN [1]

```
RP SEQUENCE FROM N.A.
RC STRAIN-IS2D;
RA Moore J.B., Beverley S.M.;
RT "Identification of a Leishmania mediating folate/methotrexate uptake by
   genetic rescue of a Leishmania transport mutant.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF084472; AAD52049.1; -.
SQ SEQUENCE 700 AA; 76420 MW; 3A3450E8D57C42AE CRC64;

Query Match          53.5%; Score 38; DB 5; Length 700;
Best Local Similarity 43.8%; Pred. No. 3.3e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 6; Gaps 1;

QY 2 PKPQQ-----WFLWM 11
   ||| |||:|
Db 194 PKPGPSMVSINWIFWM 209

RESULT 109
Q9M7X1 PRELIMINARY; PRT; 722 AA.
AC Q9M7X1.
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DE MJK13.4 PROTEIN.
GN MJK13.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome III P1 MJK13 genomic sequence.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024081; AAF35404.1; -.
SQ SEQUENCE 722 AA; 81163 MW; 7F6861ABC2AC29E9 CRC64;

Query Match          53.5%; Score 38; DB 10; Length 722;
Best Local Similarity 62.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 PQQWFWM 11
   | ||| |
Db 528 PDHWFWM 535

RESULT 110
Q9DXA2 PRELIMINARY; PRT; 901 AA.
AC Q9DXA2.
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TRENBLrel. 16, Last annotation update)
DE ENHANCIN.
OS Choristoneura fumiferana granulovirus.
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=56947;
RN [1]
RP SEQUENCE FROM N.A.
RA Bourassa A., Bergeron J., Merzouki A., Guertin C.;
RT "Enhancin of Choristoneura fumiferana granulovirus.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF319939; AAG33872.1; -.
SQ SEQUENCE 901 AA; 104208 MW; 267F182EAE85C6D8 CRC64;
```

```
Query Match          53.5%; Score 38; DB 12; Length 901;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQWFWM 10
   | || | ||
Db 353 PYPQIWSWL 361

RESULT 111
P74693 PRELIMINARY; PRT; 909 AA.
AC P74693.
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE ACRIFLAVINE RESISTANCE PROTEIN.
GN ACRB OR SLR0454.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
   Synechocystis sp. strain PCC6803. II. Sequence determination of the
   entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
DR EMBL; D90917; BAA18811.1; -.
DR InterPro; IPR001036; ACR_tran.
DR Pfam; PF00873; ACR_tran; 3.
KW Complete proteome.
SQ SEQUENCE 909 AA; 97673 MW; 96E59A0AE9FF6DFB CRC64;

Query Match          53.5%; Score 38; DB 2; Length 909;
Best Local Similarity 71.4%; Pred. No. 4.3e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQQWF 8
   | : || |
Db 445 PRPQSWF 451

RESULT 112
O00908 PRELIMINARY; PRT; 956 AA.
ID O00908
AC O00908.
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)
DE POLYTHREONINE PROTEIN (FRAGMENT).
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-GCHI;
RA Carraway M., Tzipori S., Widmer G.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U83169; AAB61362.1; -.
FT NON_TER 1
FT NON_TER 956
SQ SEQUENCE 956 AA; 100783 MW; 61613B6D7F1895C6 CRC64;
```

Query Match 53.5%; Score 38; DB 5; Length 956;  
Best Local Similarity 62.5%; Pred. No. 4.5e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
Db 250 KPDEWCWL 257

## RESULT 113

075339 PRELIMINARY; PRT; 1184 AA.  
AC O75339;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE CARTILAGE INTERMEDIATE LAYER PROTEIN.  
GN CILP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUP-ARTICULAR CARTILAGE;  
RX MEDLINE=98389785; PubMed=9722584;  
RA Lorenzo P., Neame P., Sommarin Y., Heinegard D.;  
RT "Cloning and deduced amino acid sequence of a novel cartilage protein (CILP) identifies a proform including a nucleotide pyrophosphatase."  
RT J. Biol. Chem. 273:23469-23475(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Nakamura I., Okawa A., Ikegawa S., Takaoka K., Nakamura Y.;  
RT "Genomic organization, mapping, and polymorphisms of the gene encoding human cartilage intermediate layer protein (CILP)."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Lorenzo P., Aman P., Sommarin Y., Heinegard D.;  
RT "Pro-CILP: Gene structure and chromosomal localization."  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX DOMAIN.

EMBL; AF035408; AAC33838.1; -;  
EMBL; AB022430; BAA76692.1; -;  
EMBL; AF035455; AAF14889.1; -;  
EMBL; AF035448; AAF14889.1; JOINED.  
EMBL; AF035449; AAF14889.1; JOINED.  
EMBL; AF035451; AAF14889.1; JOINED.  
EMBL; AF035453; AAF14889.1; JOINED.  
DR InterPro: IPR002086; Aldehyde dehydr.  
DR InterPro: IPR001451; Hexapep\_transf.  
DR InterPro: IPR003598; Ig\_C2.  
DR InterPro: IPR003006; Ig\_MHC.  
DR InterPro: IPR000884; TSP1.  
PFam: PF00047; Ig\_1.  
PFam: PF00090; TSP\_1.  
SMART; SM00408; IGC2; 1.  
SMART; SM00209; TSP1; 1.  
DR PROSITE; PS00070; ALDEHYDE\_DEHYDR\_CYS; UNKNOWN\_1.  
DR PROSITE; PS00101; HEXAPEP\_TRANSFERASES; UNKNOWN\_1.  
DR PROSITE; PS50092; TSP1; 1.

SEQUENCE 1184 AA; 132538 MW; 4449F05537CC99C3 CRC64;

Query Match 53.5%; Score 38; DB 4; Length 1184;  
Best Local Similarity 44.4%; Pred. No. 5.6e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKQOWFW 9  
Db 335 KRPDKYFW 343

## RESULT 114

Q9QYV8 PRELIMINARY; PRT; 1216 AA.  
ID Q9QYV8  
AC Q9QYV8;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE DNA POLYMERASE GAMMA (EC 2.7.7.7).  
GN MIP1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUP-LIVER;  
RA Sun Q., Popanda O., Thielmann H.W.;  
RT "Mitochondrial DNA polymerase gamma from Novikoff hepatoma cells differs from that of normal rat liver in cDNA sequence and kinetic properties."  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ245646; CAB56206.1; -;  
DR InterPro: IPR001098; DNA\_pol\_A.  
DR InterPro: IPR001179; FKBP\_PPIase.  
DR InterPro: IPR002297; DNA\_polG.  
DR Pfam; PF00254; FKBP; 1.  
DR Pfam; PF00476; DNA\_pol\_A; 1.  
DR PRINTS; PR00867; DNAPOLG.  
DR PROSITE; PS00447; DNA\_POLYMERASE\_A; 1.  
DR SMART; SM00482; POLAC; 1.  
KW Transferase; Nucleotidyltransferase.  
SQ SEQUENCE 1216 AA; 136855 MW; A74CD68B68DB690C CRC64;

Query Match 53.5%; Score 38; DB 11; Length 1216;  
Best Local Similarity 50.0%; Pred. No. 5.7e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KPQOWFW 9  
Db 147 PQPRKWW 154

## RESULT 115

Q9QYV7 PRELIMINARY; PRT; 1216 AA.  
ID Q9QYV7  
AC Q9QYV7;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE DNA POLYMERASE GAMMA (EC 2.7.7.7).  
GN MIP1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUP-LIVER;  
RA Sun Q., Popanda O., Thielmann H.W.;  
RT "Mitochondrial DNA polymerase gamma from Novikoff hepatoma cells differs from that of normal rat liver in cDNA sequence and kinetic properties."  
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ245647; CAB56207.1; -;  
DR InterPro: IPR001098; DNA\_pol\_A.  
DR InterPro: IPR001179; FKBP\_PPIase.  
DR InterPro: IPR002297; DNA\_polG.  
DR Pfam; PF00254; FKBP; 1.  
DR Pfam; PF00476; DNA\_pol\_A; 1.  
DR PRINTS; PR00867; DNAPOLG.

```
DR PROSITE; PS00447; DNA_POLYMERASE_A; 1.
DR SMART; SM00482; POLAC; 1.
KW Transferase; Nucleotidyltransferase.
SQ SEQUENCE 1216 AA; 136836 MW; 3E55311408373ABE CRC64;

Query Match          53.5%; Score 38; DB 11; Length 1216;
Best Local Similarity 50.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKQQWFW 9
   |:::|
Db 147 PQRKQWV 154

RESULT 116
P81137
AC P81137 PRELIMINARY; PRT; 1528 AA.
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE INSECTICIDAL TOXIN RECEPTOR BT-R1 PRECURSOR.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=MIDGUT;
RX MEDLINE=95197553; PubMed=7890666;
RA Vadlamudi R.K., Weber E., Ji I., Ji T.H., Bulla L.A. Jr.;
RT "Cloning and expression of a receptor for an insecticidal toxin of
   Bacillus thuringiensis.";
RL J. Biol. Chem. 270:5490-5494(1995).
CC -!- FUNCTION: BINDS TO THE CRYIA(B) TOXIN OF BACILLUS THURINGIENSIS
CC -!- SUBSP. BERLINER, LEADING TO THE DEATH OF M.SEXTA.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN
CC -!- SIMILARITY: RELATED TO THE CADHERIN FAMILY OF CELL ADHESION
CC MOLECULES. CONTAINS 11 CADHERIN-TYPE REPEATS.
DR InterPro; IPR001525; C5_DNA_meth.
DR Pfam; PF00028; Cadherin.
DR SMART; SM00112; CA; 10.
DR PROSITE; PS00095; C5_MTASE_2; UNKNOWN_1.
DR PROSITE; PS00232; CADHERIN_1; 1.
DR PROSITE; PS0268; CADHERIN_2; 10.
DR Receptor; Glycoprotein; Transmembrane; Signal; Repeat; Cell adhesion.
FT SIGNAL 1 21
FT CHAIN 22 1528 INSECTICIDAL TOXIN RECEPTOR BT-R1.
FT DOMAIN 22 1405 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 72 1353 11 X APPROXIMATE TANDEM REPEATS.
FT TRANSMEM 1406 1428 POTENTIAL.
FT DOMAIN 1429 1528 CYTOPLASMIC (POTENTIAL).
FT REPEAT 72 176 CADHERIN 1.
FT REPEAT 177 289 CADHERIN 2.
FT REPEAT 290 397 CADHERIN 3.
FT REPEAT 398 500 CADHERIN 4.
FT REPEAT 501 623 CADHERIN 5.
FT REPEAT 624 757 CADHERIN 6.
FT REPEAT 758 882 CADHERIN 7.
FT REPEAT 883 1004 CADHERIN 8.
FT REPEAT 1005 1121 CADHERIN 9.
FT REPEAT 1122 1242 CADHERIN 10.
FT REPEAT 1243 1353 CADHERIN 11.
FT CARBOHYD 127 120 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 305 305 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 575 575 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT CARBOHYD 709 709 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 865 865 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 894 894 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 986 986 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1049 1049 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1069 1069 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1104 1104 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1528 AA; 171986 MW; CF678E01D700C91D CRC64;

Query Match          53.5%; Score 38; DB 5; Length 1528;
Best Local Similarity 71.4%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQQWFWL 10
   |:::|
Db 233 PNQWFWL 239

RESULT 117
Q9GPJ9
ID Q9GPJ9 PRELIMINARY; PRT; 1717 AA.
AC Q9GPJ9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CADHERIN-RELATED PROTEIN RECEPTOR BT-R1.
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC Sphingioidea; Sphingidae; Sphinginae; Manduca.
OX NCBI_TaxID=7130;
RN [1]
RP SEQUENCE FROM N.A.
RA Dorsch J.A., Maaty W.S.A., Griko N.B., Candias M., Bulla L.A. Jr.;
RT "A Cadherin-related Protein Receptor, BT-R1, in the Midgut Epithelium
   of Manduca sexta Mediates Toxicity for Bacillus thuringiensis CryIa
   Toxins.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE CADHERIN FAMILY.
DR EMBL; AF319973; AAG37912.1; -.
DR InterPro; IPR001525; C5_DNA_meth.
DR InterPro; IPR002126; Cadherin.
DR Pfam; PF00028; cadherin; 6.
DR SMART; SM00112; CA; 10.
DR PROSITE; PS00095; C5_MTASE_2; UNKNOWN_1.
DR PROSITE; PS00232; CADHERIN_1; 1.
DR PROSITE; PS0268; CADHERIN_2; 10.
DR Receptor; Glycoprotein; Cell adhesion; Glycoprotein; Receptor.
KW Calcium-binding; Cell adhesion; Glycoprotein; Receptor.
SQ SEQUENCE 1717 AA; 192305 MW; FEC4A48B098B817E CRC64;

Query Match          53.5%; Score 38; DB 5; Length 1717;
Best Local Similarity 71.4%; Pred. No. 8e+02;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQQWFWL 10
   |:::|
Db 234 PNQWFWL 240

RESULT 118
Q96503
ID Q96503 PRELIMINARY; PRT; 1832 AA.
AC Q96503;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE GP900.
OS Cryptosporidium parvum.
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;
OC Cryptosporidiidae; Cryptosporidium.
OX NCBI_TaxID=5807;
```

RA SEQUENCE FROM N.A.  
RX MEDLINE-99066935; PubMed=9851610;  
RA Barnes D.A., Bonnin A., Huang J.X., Gousset L., Wu J., Gut J.,  
RA Doyle P., Dubremetz J.F., Ward H., Petersen C.;  
RT "A novel multi-domain mucin-like glycoprotein of Cryptosporidium  
RT parvum mediates invasion.";  
RL Mol. Biochem. Parasitol. 96:93-110(1998).  
DR EMBL: AF068065; AAC98153.1; -;  
SQ SEQUENCE 1832 AA; 192653 MW; 590E6ACB16BBE0D2 CRC64;

Query Match 53.5%; Score 38; DB 5; Length 1832;  
Best Local Similarity 62.5%; Pred. No. 8.5e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQWFNL 10  
|| : ||  
Db 527 KPDEWCWL 534

RESULT 119  
Q9VI22 PRELIMINARY; PRT; 2237 AA.  
ID Q9VI22  
AC Q9VI22;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE CG10272 PROTEIN.  
GN CG10272.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE-20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Flossler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003675; AAF54122.1; -;  
DR FlyBase: FBgn0037444; AA\_trna\_ligase\_II, CGI0272.  
DR InterPro: IPR002106; AA\_trna\_ligase\_II-2; UNKNOWN\_1.  
DR PROSITE: PS00339; AA\_trna\_ligase\_II-2; UNKNOWN\_1.  
SQ SEQUENCE 2237 AA; 242406 MW; 2E4A397306BF6E57 CRC64;

Query Match 53.5%; Score 38; DB 5; Length 2237;  
Best Local Similarity 54.5%; Pred. No. 1e+03;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPOQWFNL 11  
|| : ||  
Db 2199 RPKGKDWDMWS 2209

RESULT 120  
Q99IT0 PRELIMINARY; PRT; 100 AA.  
ID Q99IT0  
AC Q99IT0;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE HYPOTHETICAL 11.5 KDA PROTEIN.  
OS unclassified; environmental samples.  
OC NCBI\_TaxID=155900;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Stokes H.W., Nield B.S., Mabbutt B.C., Nevalainen H., Holmes A.J.,  
RA Gillings M.R.;  
RT "Novel and diverse integron-like gene cassettes are prevalent in  
RT natural environments.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF349104; AAK28610.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 100 AA; 11451 MW; A39185CC7D9D674A CRC64;

Query Match 52.8%; Score 37.5; DB 14; Length 100;  
Best Local Similarity 58.3%; Pred. No. 60;  
Matches 7; Conservative 1; Mismatches 1; Indels 3; Gaps 1;

QY 2 PKPQOWF---WL 10  
|| : ||  
Db 48 PKPQOWSYRWL 59

RESULT 121  
Q99ZD5 PRELIMINARY; PRT; 213 AA.  
ID Q99ZD5  
AC Q99ZD5;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE PUTATIVE AMINO ACID ABC TRANSPORTER (PERMEASE PROTEIN).  
GN SPY1276.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
RN NCBI\_TaxID=1314;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-SF370;  
RX MEDLINE-21192684; PubMed=11296296;  
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K.,  
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;

RT \*Complete genome sequence of an M1 strain of Streptococcus pyogenes.\*;  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
DR EMBL; AE006566; AAK34126.1; -  
KW Complete proteome.  
SQ SEQUENCE 213 AA; 23990 MW; FCAB7D5BA2E32097 CRC64;

Query Match 52.8%; Score 37.5; DB 2; Length 213;  
Best Local Similarity 50.08; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 5; Gaps 1;

Qy 3 RPQWFM-----WLM 11  
||| ||| |||  
Db 45 KPLQWFLTYVWM 58

## RESULT 122

O9CY68 PRELIMINARY; PRT; 300 AA.  
ID AC Q9CY68  
AC Q9CY68; 17, Created  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DE 7 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
DE CLONE:C430041K09, FULL INSERT SEQUENCE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzairelli J., Sakamoto N.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
RA Hayashizaki Y.,  
RT \*Functional annotation of a full-length mouse cDNA collection.\*;  
RL Nature 409:685-690(2001).  
DR EMBL; AK021249; BAB32347.1; -  
SQ SEQUENCE 300 AA; 34815 MW; 94D65F47AD4A208D CRC64;

Query Match 52.8%; Score 37.5; DB 11; Length 300;  
Best Local Similarity 63.6%; Pred. No. 1.8e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 RPQPOW-FWL 10  
||| ||| |||  
Db 97 RHPAYWFWL 107

## RESULT 123

O921X2 PRELIMINARY; PRT; 473 AA.  
ID AC Q921X2  
AC Q921X2; 10, Created  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)

DE PHOSPHATIDYL SERINE SYNTHASE-2.  
GN PTDSS2  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=LIVER;  
RA Stone S.J., Vance J.E.;  
RT \*Phosphatidylserine synthase-2 from mouse liver.\*;  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF09053; AAC98383.1; -  
DR MGD; MGI:1351664; Ptdss2.  
SQ SEQUENCE 473 AA; 55004 MW; 66402CE1EB7D0B07 CRC64;

Query Match 52.8%; Score 37.5; DB 11; Length 473;  
Best Local Similarity 63.6%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 RPQPOW-FWL 10  
||| ||| |||  
Db 97 RHPAYWFWL 107

## RESULT 124

O08888 PRELIMINARY; PRT; 474 AA.  
ID AC O08888  
AC O08888; 04, Created  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)  
DT 01-AUG-1998 (TREMBlrel. 07, Last annotation update)  
DE PHOSPHATIDYL SERINE SYNTHASE II.  
GN PSSB.  
OS Cricetus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetus.  
OX NCBI\_TaxID=10029;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kuge O., Saito K., Masahiro N.;  
RL J. Biol. Chem. 0:0-0(0).  
DR EMBL; AB004109; BAA20355.1; -  
SQ SEQUENCE 474 AA; 55004 MW; 81942EA310F538C2 CRC64;

Query Match 52.8%; Score 37.5; DB 11; Length 474;  
Best Local Similarity 63.6%; Pred. No. 2.7e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 1 RPQPOW-FWL 10  
||| ||| |||  
Db 97 RHPAYWFWL 107

## RESULT 125

O9BVG9 PRELIMINARY; PRT; 487 AA.  
ID AC O9BVG9  
AC O9BVG9; 17, Created  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE SIMILAR TO PHOSPHATIDYL SERINE SYNTHASE 2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=NEUROBLASTOMA;  
RA Strausberg R.;

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC001210; AA01210.1; -  
SQ SEQUENCE 487 AA; 56253 MW; E02508F894841A4F CRC64;

Query Match 52.8%; Score 37.5; DB 4; Length 487;  
Best Local Similarity 63.6%; Pred. No. 2.8e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 RPKPOQW-FWL 10  
||| |||||  
Db 119 RHPAYWRFWL 129

RESULT 126

ID O92611 PRELIMINARY; PRT; 1177 AA.  
AC O92611;  
DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE DNA-BINDING PROTEIN.  
GN DBP.  
OS Pseudorabies virus.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=10345;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=TNL;  
RX MEDLINE=98455382; PubMed=9784061;  
RA Wu S.-L., Hsiang C.-Y., Ho T.-Y., Chang T.-J.;  
RT "Identification, expression, and characterization of the pseudorabies  
RT virus DNA-binding protein gene and gene product.";  
RL Virus Res. 56:1-9(1998).  
DR EMBL: U80909; AAC63429.1; -  
DR InterPro: IPR000635; Viral\_DNA\_bind.  
DR Pfam: PF00747; Viral\_DNA\_bp; 1.  
KW DNA-binding.  
SQ SEQUENCE 1177 AA; 125408 MW; BA87AF9CFC961707 CRC64;

Query Match 52.8%; Score 37.5; DB 12; Length 1177;  
Best Local Similarity 60.0%; Pred. No. 6.6e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PKPOQWFWM 11  
||| |||||  
Db 820 PNP-QWFWM 828

RESULT 127

ID Q9E1Y7 PRELIMINARY; PRT; 1194 AA.  
AC Q9E1Y7;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE SSDNA BINDING PROTEIN.  
OS Carcipothecine herpesvirus 7.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=35245;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Gray W.L., Starnes H.B., White M.W., Ashburn C.V., Mahalingam R.;  
RT "Complete Sequence of the Simian Varicella Virus Genome."  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF275348; AAG27202.1; -  
DR InterPro: IPR000635; Viral\_DNA\_bind.  
DR Pfam: PF00747; Viral\_DNA\_bp; 1.  
SQ SEQUENCE 1194 AA; 131968 MW; EBA7F3C841965897 CRC64;

Query Match 52.8%; Score 37.5; DB 12; Length 1194;  
Best Local Similarity 60.0%; Pred. No. 6.7e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PKPOQWFWM 11  
||| |||||  
Db 830 PNP-QWFWM 838

RESULT 128

ID Q69101 PRELIMINARY; PRT; 1197 AA.  
AC Q69101;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE DNA BINDING PROTEIN ICP8.  
OS Herpes simplex virus (type 2).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10310;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KN;  
RX MEDLINE=93228441; PubMed=8385914;  
RA Toh Y., Tanaka S., Liu Y., Mori R.;  
RT "Nucleotide sequence of the major DNA-binding protein gene of herpes  
RT simplex virus type 2 and a comparison with the type 1.";  
RL Arch. Virol. 129:183-196(1993).  
DR EMBL: D10658; BAA01507.1; -  
DR InterPro: IPR000635; Viral\_DNA\_bind.  
DR Pfam: PF00747; Viral\_DNA\_bp; 1.  
SQ SEQUENCE 1197 AA; 128470 MW; AA3ADA75B8865BFE CRC64;

Query Match 52.8%; Score 37.5; DB 12; Length 1197;  
Best Local Similarity 66.7%; Pred. No. 6.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPKPOQFW 9  
||| |||||  
Db 837 QPNP-QFW 844

RESULT 129

ID Q89549 PRELIMINARY; PRT; 1203 AA.  
AC Q89549;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE UL29.  
GN UL29.  
OS Bovine herpesvirus 1.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Varicellovirus.  
OX NCBI\_TaxID=10320;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COOPER;  
RA Schwyzner M., Vilek C., Lowery D.E., Bello L.J., Meyer G., Misra V.,  
RA Thiry E., Paces V.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COOPER;  
RA Meyer G., Vilek C., Paces V., Pastoret P., Thiry E., Schwyzner M.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JURA.  
RA Schwyzner M., Vilek C., Lowery D.E., Bello L.J., Meyer G., Misra V.;



RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JURA;  
 RA Schwyzer M.;  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Z78205; CAB01596.1; -;  
 DR EMBL; X94677; CAA64336.1; -;  
 DR EMBL; AJ004801; CAA06104.1; -;  
 DR InterPro: IPR000635; Viral\_DNA\_bind.  
 DR Pfam; PF00747; Viral\_DNA\_dp; 1.  
 SQ SEQUENCE 1203 AA; 127409 MW; 8299D64966A9654F CRC64;

Query Match 52.8%; Score 37.5; DB 12; Length 1203;  
 Best Local Similarity 60.0%; Pred. No. 6.8e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 2 PKPQOWFLM 11  
 |||||  
 Db 840 PNP-QWFWTL 848

## RESULT 130

O39273  
 ID O39273 PRELIMINARY; PRT; 1208 AA.  
 AC O39273;  
 DT 01-JAN-1998 (TRENBLrel. 05, Created)  
 DT 01-JUN-1998 (TRENBLrel. 05, Last sequence update)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
 DE COUNTERPART OF HSV-1 GENE UL29 AND VZV GENE 29.  
 GN 31.  
 OS Equine herpesvirus 4.  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 CC Alphaherpesvirinae; Varicellovirus.  
 OX NCBI\_TaxID=10331;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NS80567;  
 RX MEDLINE=98264497; PubMed=9603335;  
 RA Telford E.A.R., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;  
 RT "The DNA sequence of equine herpesvirus-4.";  
 RL J. Gen. Virol. 79:1197-1203(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NS80567;  
 RA Telford E.A., Watson M.S., Perry J., Cullinane A.A., Davison A.J.;  
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF030027; AAC59547.1; -;  
 DR InterPro: IPR000635; Viral\_DNA\_bind.  
 DR Pfam; PF00747; Viral\_DNA\_dp; 1.  
 SQ SEQUENCE 1208 AA; 130607 MW; 17699FBD9238C4CB CRC64;

Query Match 52.8%; Score 37.5; DB 12; Length 1208;  
 Best Local Similarity 60.0%; Pred. No. 6.8e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

Qy 2 PKPQOWFLM 11  
 |||||  
 Db 844 PNP-QWFWTL 852

## RESULT 131

O9TBB3  
 ID O9TBB3 PRELIMINARY; PRT; 53 AA.  
 AC O9TBB3;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DE ATPASE 8 (FRAGMENT).  
 OS Tadorna variegata.  
 OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Tadorna.  
 OX NCBI\_TaxID=107024;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sorenson M.D., Cooper A., Paxinos E.E., Quinn T.W., James H.F.,  
 RA Olson S.L., Fleischer R.C.;  
 RT "Relationships of the extinct moa-nalos, flightless Hawaiian waterfowl,  
 RT based on ancient DNA.";  
 RL Proc. R. Soc. Lond. B, Biol. Sci. 0:0-0(1999).  
 DR EMBL; AF173743; AAF07025.1; -;  
 DR InterPro: IPR001421; ATP-synt\_8.  
 DR InterPro: IPR003237; Avian\_mito\_ATPase\_8.  
 DR Pfam; PF00895; ATP-synt\_8; 1.  
 DR ProDom; PD161863; Avian\_mito\_ATPase\_8; 1.  
 KW Mitochondrion.  
 FT NON\_TER 53  
 SQ SEQUENCE 53 AA; 6044 MW; B6AC0B525B992891 CRC64;

Query Match 52.1%; Score 37; DB 8; Length 53;  
 Best Local Similarity 62.5%; Pred. No. 39;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PKPQOWFW 9  
 |||||  
 Db 45 PKPTPAW 52

## RESULT 132

O9TBB2  
 ID O9TBB2 PRELIMINARY; PRT; 53 AA.  
 AC O9TBB2;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
 DE ATPASE 8 (FRAGMENT).  
 OS Tadorna tadorna.  
 OG Mitochondrion.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Anseriformes; Anatidae; Tadorna.  
 OX NCBI\_TaxID=75865;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sorenson M.D., Cooper A., Paxinos E.E., Quinn T.W., James H.F.,  
 RA Olson S.L., Fleischer R.C.;  
 RT "Relationships of the extinct moa-nalos, flightless Hawaiian waterfowl,  
 RT based on ancient DNA.";  
 RL Proc. R. Soc. Lond. B, Biol. Sci. 0:0-0(1999).  
 DR EMBL; AF173744; AAF07026.1; -;  
 DR InterPro: IPR001421; ATP-synt\_8.  
 DR InterPro: IPR003237; Avian\_mito\_ATPase\_8.  
 DR Pfam; PF00895; ATP-synt\_8; 1.  
 DR ProDom; PD161863; Avian\_mito\_ATPase\_8; 1.  
 KW Mitochondrion.  
 FT NON\_TER 53  
 SQ SEQUENCE 53 AA; 5970 MW; 4ELE99525B855882 CRC64;

Query Match 52.1%; Score 37; DB 8; Length 53;  
 Best Local Similarity 62.5%; Pred. No. 39;  
 Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PKPQOWFW 9  
 |||||  
 Db 45 PKPAPWAW 52

## RESULT 133

O9RM36  
 ID O9RM36 PRELIMINARY; PRT; 116 AA.  
 AC O9RM36;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE HYPOTHETICAL PROTEIN (FRAGMENT).  
OS Escherichia coli.  
OG Plasmid por17.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC PLASMID-POR17;  
RX MEDLINE=84067491; PubMed=6358799;  
RA Tait R.C., Kado C.I., Rodriguez R.L.;  
RT "A comparison of the origin of replication of pSA with R6K.";  
RL Mol. Gen. Genet. 192:32-38(1983).  
DR EMBL: X00060; CAB56196.1; -;  
KW Plasmid.  
FT NON\_TER 1 1  
FT NON\_TER 116 116  
SQ SEQUENCE 116 AA; E41EBB9C4F154A52 CRC64;

Query Match 52.1%; Score 37; DB 2; Length 116;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WFWLM 11  
| | | | |  
DB 103 WFWLM 107

## RESULT 134

ID Q9ADB8 PRELIMINARY; PRT; 181 AA.  
AC Q9ADB8;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DE HYPOTHETICAL 20.5 KDA PROTEIN.  
GN SCBAC5H2.19.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Saunders D.C., Harris D.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL: AL589707; CAC33914.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 181 AA; 20454 MW; F7C3F5048309A3C5 CRC64;

Query Match 52.1%; Score 37; DB 2; Length 181;  
Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQWFWL 10

DB 142 RPLPDLWPWL 151  
| | | | |

## RESULT 135

ID O14264 PRELIMINARY; PRT; 220 AA.  
AC O14264;  
DT 01-JAN-1998 (TReMBLrel. 05, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE HYPOTHETICAL 25.8 KDA PROTEIN C7D4.09C IN CHROMOSOME I.  
GN SPAC7D4.09C.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA Gentles S., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Z99532; CAB16726.2; -;  
DR InterPro: IPR001104; S5A\_redtsc\_C.  
KW Hypothetical protein.  
SQ SEQUENCE 220 AA; 25760 MW; 8314536BD00595C8 CRC64;

Query Match 52.1%; Score 37; DB 3; Length 220;  
Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 PQQWFW 9  
| : | | | |  
DB 61 PKRWFW 66

## RESULT 136

ID Q9V424 PRELIMINARY; PRT; 226 AA.  
AC Q9V424;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)  
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE METHYL-CPG-BINDING-DOMAIN-LIKE-PROTEIN.  
GN METHYL-CPG-BINDING-DOMAIN-LIKE-PROTEIN OR CG8208.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,



SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heilmann T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pauley J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Furl V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen R.-F., Zaveri F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RA "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003683; AAF54400.1;  
DR Flybase; FBgn0027930; MBD.  
DR InterPro; IPR001739; MBD.  
DR SMART; SM00391; MBD; 1.  
SQ SEQUENCE 314 AA; 33803 MW; 7B6BDFE63873D230 CRC64;

Query Match 52.1%; Score 37; DB 5; Length 314;  
Best Local Similarity 66.7%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQOWFW.9  
Db 198 REKPKQLFW 206

RESULT 140  
Q00853  
ID Q00893 PRELIMINARY; PRT; 331 AA.  
AC Q00893;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DR 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE PECTATE LYASE (FRAGMENT).  
GN PEL.  
OS Colletotrichum gloeosporioides (Anthracnose fungus) (Glomerella  
OS eugulata).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;  
OC Glomerella.  
OX NCBI\_TaxID=5457;  
RN [1]

SEQUENCE FROM N.A.  
RC STRAIN-CG-14;  
RA Wattad C., Dinooor A., Prusky D.;  
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U32942; AAA75471.1; -;  
DR HSP; P11073; IAIR.  
DR Mendel; 20821; Colgl; 3108; 20821.  
DR InterPro; IPR002022; Amb\_allergen.  
DR Pfam; PF00544; pec\_lyase; 1.  
DR PRINTS; PR00807; AMBALLERGEN.  
KW Lyase.  
FT NON\_TER 1 I  
SQ SEQUENCE 331 AA; 35416 MW; 092AADCEBF6DF7EB CRC64;

Query Match 52.1%; Score 37; DB 3; Length 331;  
Best Local Similarity 62.5%; Pred. No. 2.3e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQOWF 8  
Db 290 RPSPSKWF 297

RESULT 141

Q9Y231  
ID Q9V231 PRELIMINARY; PRT; 359 AA.  
AC Q9V231;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE ALPHA-3-FUCOSYLTRANSFERASE.  
GN FUT9.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-WHOLE EMBRYO;  
RA Cailliau A., Coullin P., Candellier J.J., Balanzino L., Oriol R.,  
RA Mollicone R.;  
RT "Cloning, expression and chromosome localization of a human embryonic  
RT FUT9 transcript.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9932063; PubMed=10405152;  
RA Kaneko M., Kudo T., Iwasaki H., Ikehara Y., Nishihara S., Nakagawa S.,  
RA Sasaki K., Shina T., Inoko H., Saitou N., Narimatsu H.;  
RT "Human hepatitis B virus mutants: significance of molecular changes.";  
RL FEBS Lett. 453:237-242(1999).  
DR EMBL; AJ238701; CAB41890.1; -;  
DR EMBL; AB023021; BAA81685.1; -;  
DR InterPro; IPR001503; Glyco\_transf\_10.  
DR Pfam; PF00852; Glyco\_transf\_10; 1.  
KW Transferase; Glycosyltransferase.  
SQ SEQUENCE 359 AA; 42041 MW; C90CF5C02CB644D9 CRC64;

Query Match 52.1%; Score 37; DB 4; Length 353;  
Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPQOWFWL 10  
Db 125 RPPFKWIWM 134

RESULT 142  
O88819  
ID O88819 PRELIMINARY; PRT; 359 AA.  
AC O88819;

01-NOV-1998 (TRENBLrel. 08, Created)  
01-NOV-1998 (TRENBLrel. 08, Last sequence update)  
01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
ALPHA-(1,3)-FUCOSYLTRANSFERASE (EC 2.4.1.-) (GALACTOSIDE 3-L-  
FUCOSYLTRANSFERASE) (FUCOSYLTRANSFERASE 9) (FUCT-IX) (MFUC-TIX).  
FUT9.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.  
RC STRAIN=BALB/C; TISSUE=BRAIN;  
RX MEDLINE=98434588; PubMed=9756916;  
RA Kudo T., Ikehara Y., Togayachi A., Kaneko M., Hiraga T., Sasaki K.,  
RA Narimatsu H.;  
"Expression cloning and characterization of a novel murine alpha1, 3-  
fucosyltransferase, mfuc-TIX, that synthesizes the Lewis x (CD15)  
epitope in brain and kidney.";  
J. Biol. Chem. 273:26729-26738(1998).  
CC -!- FUNCTION: MAY CATALYZE ALPHA-1,3 GLYCOSIDIC LINKAGES INVOLVED IN  
CC THE EXPRESSION OF LEWIS X AND LEWIS Y.  
CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + 1,4-BETA-D-GALACTOSYL-N-  
ACETYL-D-GLUCOSAMINYL-R = GDP + 1,4-BETA-D-GALACTOSYL-(ALPHA-  
1,3-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINYL-R.  
CC -!- PATHWAY: GLYCOSYLATION.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
CC FORM IN TRANS CISTERNAE OF GOLGI.  
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BRAIN AND KIDNEY. IT  
CC ALSO EXPRESSED IN THE STOMACH, COLON, UTERUS AND EPIDIDYMIS. NOT  
CC FOUND IN THYMUS, LIVER, SPLEEN, Ovary, LUNG, HEART, TESTIS AND  
CC SMALL INTESTINE.  
DR EMBL; AB015426; BAA33522.1; -.  
DR MGD; MGI:1330859; Fut9.  
DR InterPro; IPR001503; Glyco\_transf\_10.  
DR Pfam; PF00852; Glyco\_transf\_10; 1.  
KW Transferase; Glycosyltransferase; Transmembrane; Glycoprotein;  
KW Signal-anchor; Golgi stack.  
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 12 30 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
(POTENTIAL).  
FT DOMAIN 31 359 LUMENAL, CATALYTIC (POTENTIAL).  
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 359 AA; 42041 MW; 96A2394547F2A44E CRC64;  
Query Match 52.1%; Score 37; DB 11; Length 359;  
Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 RPKPQWFNL 10  
Db 125 RPPFKQWIM 134  
RESULT 143  
Q9JIG1  
ID Q9JIG1 PRELIMINARY; PRT; 359 AA.  
AC Q9JIG1  
DT 01-OCT-2000 (TRENBLrel. 15, Created)  
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE ALPHA-(1,3)-FUCOSYLTRANSFERASE (EC 2.4.1.-) (GALACTOSIDE 3-L-  
FUCOSYLTRANSFERASE) (FUCOSYLTRANSFERASE 9) (FUCT-IX).  
GN FUT9.  
OS Cricetus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetus.  
OX NCBI\_TaxID=10029;  
RN [1]

SEQUENCE FROM N.A.  
RP TISSUE=OVARY;  
RX MEDLINE=20166953; PubMed=10700388;  
RA Patnaik S.K., Zhang A., Shi S., Stanley P.;  
RT "Alpha(1,3)fucosyltransferases expressed by the gain-of-function  
Chinese hamster ovary glycosylation mutants LEC12, LEC29, and LEC30.";  
Arch. Biochem. Biophys. 375:322-332(2000).  
CC -!- FUNCTION: MAY CATALYZE ALPHA-1,3 GLYCOSIDIC LINKAGES INVOLVED IN  
CC THE EXPRESSION OF LEWIS X AND LEWIS Y.  
CC -!- CATALYTIC ACTIVITY: GDP-L-FUCOSE + 1,4-BETA-D-GALACTOSYL-N-  
ACETYL-D-GLUCOSAMINYL-R = GDP + 1,4-BETA-D-GALACTOSYL-(ALPHA-  
1,3-L-FUCOSYL)-N-ACETYL-D-GLUCOSAMINYL-R.  
CC -!- PATHWAY: GLYCOSYLATION.  
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND  
CC FORM IN TRANS CISTERNAE OF GOLGI.  
DR EMBL; AF230460; AAF82412.1; -.  
DR InterPro; IPR001503; Glyco\_transf\_10.  
DR Pfam; PF00852; Glyco\_transf\_10; 1.  
KW Transferase; Glycosyltransferase; Transmembrane; Glycoprotein;  
KW Signal-anchor; Golgi stack.  
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 12 30 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
(POTENTIAL).  
FT DOMAIN 31 359 LUMENAL, CATALYTIC (POTENTIAL).  
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 359 AA; 42071 MW; 9D5CD8BFF07EA902 CRC64;  
Query Match 52.1%; Score 37; DB 11; Length 359;  
Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 RPKPQWFNL 10  
Db 125 RPPFKQWIM 134  
RESULT 144  
Q99JB3  
ID Q99JB3 PRELIMINARY; PRT; 359 AA.  
AC Q99JB3  
DT 01-JUN-2001 (TRENBLrel. 17, Created)  
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE ALPHA1,3-FUCOSYLTRANSFERASE IX.  
GN FUT9.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shimoda Y., Tajima Y., Osanai T., Katsume A., Kohara M., Kudo T.,  
RA Narimatsu H., Osumi N., Sanai Y.;  
RT "Expression of Lewis x epitope in embryonic forebrain by regulating  
alpha1,3-fucosyltransferase IX expression.";  
Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RX MEDLINE=20472964; PubMed=11020213;  
RA Baboval T., Henion T., Kinnally E., Smith F.I.;  
RT "Molecular cloning of rat alpha1,3-fucosyltransferase IX (Fuc-TIX) and  
comparison of the expression of fuc-TIV and fuc-TIX genes during rat  
postnatal cerebellum development.";  
J. Neurosci. Res. 62:206-215(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RA Smith F.I., Baboval T.;  
Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN [1]

```
DR EMBL; AB049819; BAB40953.1; -.
DR EMBL; AF345993; AAK16591.1; -.
KW Transferase; Glycosyltransferase.
SQ SEQUENCE 359 AA; 42037 MW; 369B4A7BD0C6CC80 CRC64;

Query Match 52.1%; Score 37; DB 11; Length 359;
Best Local Similarity 50.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 10
||| | | |
DB 125 RPPQKRWI 134

RESULT 145
Q9TQ03 PRELIMINARY; PRT; 365 AA.
AC Q9TQ03;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ALPHA-(1,3)-FUCOSYLTRANSFERASE (EC 2.4.1.-).
GN FUTB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN, COLON, HEART, LIVER, LUNG, TESTIS, SPLEEN, AND KIDNEY;
RX MEDLINE=20022985; PubMed=10555285;
RA Wierlinckx A., Mercier D., Oulmouden A., Petit J.M., Julien R.;
RT "Complete genomic organization of futb encoding a bovine alpha3-
RT fucosyltransferase: exons in human orthologous genes emerged from
RT ancestral intronic sequences."
RL Mol. Biol. Evol. 16:1535-1547(1999).
CC -1- FUNCTION: MAY CATALYZE ALPHA-1,3 GLYCOSIDIC LINKAGES INVOLVED IN
CC THE EXPRESSION OF LEWIS X/SSEA-1 AND VIM-2 ANTIGENS.
CC -1- PATHWAY: GLYCOSYLATION.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
DR EMBL; AJ132776; CAA10775.1; -.
DR EMBL; AJ132773; CAA10772.1; -.
DR EMBL; AJ132774; CAA10773.1; -.
DR EMBL; AJ132775; CAA10774.1; -.
DR EMBL; AJ132772; CAA10771.1; -.
DR InterPro; IPR001503; Glyco_transf_10.
DR Pfam; PF00852; Glyco_transf_10; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Glycoprotein;
KW Signal-anchor; Golgi stack.
FT DOMAIN 1 15 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 16 34 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 35 365 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 100 100 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 111 111 F -> L (IN CAA10771).
FT CONFLICT 122 122 K -> Q (IN CAA10771).
FT CONFLICT 132 133 AD -> PG (IN CAA10771).
SQ SEQUENCE 365 AA; 42720 MW; 0E5F5F8002AF5A8D CRC64;

Query Match 52.1%; Score 37; DB 6; Length 365;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9
||| | | |
DB 130 RPAQQRWV 138

us-09-988-792-2.50pct.rspt
```

```
RESULT 146
Q9KQW4 PRELIMINARY; PRT; 406 AA.
AC Q9KQW4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN VC1884.
GN VC1884.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AE004263; AAF95032.1; -.
DR TIGR; VC1884; -.
DR InterPro; IPR003838; DUF214.
DR Pfam; PF02687; DUF214; 1.
KW Complete proteome.
SQ SEQUENCE 406 AA; 43747 MW; 4879D14D30442588 CRC64;

Query Match 52.1%; Score 37; DB 2; Length 406;
Best Local Similarity 55.6%; Pred. No. 2.8e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9
||| | | |
DB 243 QPLPQDQW 251

us-09-988-792-2.50pct.rspt
```

"A set of ordered cosmids and a detailed genetic and physical map for

RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL; AL445983; CAC14362.1; -.  
DR InterPro; IPR001958; TCR\_beta.  
DR PRINTS; PR01035; TCRETA.  
SQ SEQUENCE 421 AA; 43821 MW; A0B747AFF09EA6FF CRC64;

Query Match 52.1%; Score 37; DB 2; Length 421;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WFWLM 11  
|||  
DB 372 WFWLM 376

RESULT 148  
Q9CK08 PRELIMINARY; PRT; 467 AA.  
AC Q9CK08;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE HYPOTHETICAL PROTEIN PM1829.  
GN PM1829.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PM70;  
RX MEDLINE-21145866; PubMed=11248100;  
RA May B.-J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
DR EMBL; AE006220; AAK03913.1; -.  
DR InterPro; IPR002035; VWFA.  
DR SMART; SM00327; VWA; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 467 AA; 52878 MW; 6AF5BDABCC75A776 CRC64;

Query Match 52.1%; Score 37; DB 2; Length 467;  
Best Local Similarity 57.1%; Pred. No. 3.2e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPQOWFW 9  
|| :||  
DB 156 KPTRWY 162

RESULT 149  
Q9FDI3 PRELIMINARY; PRT; 529 AA.  
AC Q9FDI3;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE CYCLOHEXANONE MONOOXYGENASE 2.  
GN CHNB2.  
OS Brevibacterium sp. HCU.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Micrococciaceae; Brevibacteriaceae; Brevibacterium.  
OX NCBI\_TaxID=133406;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HCU;  
RX MEDLINE-20353458; PubMed=10894733;  
RA Brzostowicz P.C., Gibson K.L., Thomas S.M., Blasko M.S.,  
RA Rouviere P.E.;  
RT "Simultaneous Identification of Two Cyclohexanone Oxidation Genes from

RT an Environmental Brevibacterium Isolate Using mRNA Differential  
RT Display.";  
RL J. Bacteriol. 182:4241-4248(2000).  
CC -I- COFACTOR: FAD (BY SIMILARITY).  
DR EMBL; AF257215; AAG01290.1; -.  
DR InterPro; IPR001327; FAD\_pyr\_redox.  
DR InterPro; IPR001100; pyr\_redox.  
DR PRINTS; PR00368; FADPNR.  
DR PRINTS; PR00411; PNDRDTASEI.  
KW FAD: Flavoprotein; Monooxygenase; Oxidoreductase.  
SQ SEQUENCE 529 AA; 59143 MW; 71DE09C8A4E441BF CRC64;

Query Match 52.1%; Score 37; DB 2; Length 529;  
Best Local Similarity 50.0%; Pred. No. 3.6e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPOQWFW 9  
|| :||  
DB 501 PKAKSNYW 508

RESULT 150  
Q9KLT3 PRELIMINARY; PRT; 536 AA.  
AC Q9KLT3;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE METHYL-ACCEPTING CHEMOTAXIS PROTEIN.  
GN VCA0658.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
RX MEDLINE-20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
cholerae.";  
RL Nature 406:477-483(2000).  
DR EMBL; AE004395; AAF96559.1; -.  
DR TIGR; VCA0658; -.  
DR InterPro; IPR000122; Chemotaxis\_transducer.  
DR InterPro; IPR001610; PAC.  
DR InterPro; IPR000014; PAS.  
DR Pfam; PF00015; MCPsignal; 1.  
DR Pfam; PF00785; PAC; 1.  
DR Pfam; PF00989; PAC; 1.  
DR SMART; SM00283; NA; 1.  
DR SMART; SM00086; PAC; 1.  
DR SMART; SM00091; PAC; 1.  
KW Complete proteome.  
SQ SEQUENCE 536 AA; 59005 MW; C56D363FDD020CA45 CRC64;

Query Match 52.1%; Score 37; DB 2; Length 536;  
Best Local Similarity 71.4%; Pred. No. 3.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQOWFWL 10  
|:| |  
DB 188 PQOWWL 194

RESULT 151





RT "Identification of the fourth member of the mammalian endoprotease  
RT family homologous to the yeast Kex2 protease. Its testis-specific  
RT expression.";  
RL J. Biol. Chem. 267:5897-5900(1992).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=C57BL/6; TISSUE=LIVER;  
RC MEDLINE=93078790; PubMed=1448111;  
RA Seidah N.G., Day R., Hamelin J., Gaspar A., Collard M.W., Chretien M.;  
RT "Testicular expression of PC4 in the rat: molecular diversity of a  
RT novel germ cell-specific kex2/subtilisin-like proprotein convertase.";  
RT Mol. Endocrinol. 6:1559-1570(1993).  
RN [3]  
RN SEQUENCE FROM N.A.  
RP STRAIN=C57BL/6; TISSUE=LIVER;  
RC MEDLINE=94292203; PubMed=8020970;  
RA Mbikay M., Raffin-Sanson M.L., Tadros H., Sirosis F., Seidah N.G.,  
RA Chretien M.;  
RT "Structure of the gene for the testis-specific proprotein convertase 4  
RT and of its alternate messenger RNA isoforms.";  
RL Genomics 20:231-237(1994).  
DR EMBL; L21221; AAA39973.1; JOINED.  
DR EMBL; L21210; AAA39973.1; JOINED.  
DR EMBL; L21211; AAA39973.1; JOINED.  
DR EMBL; L21212; AAA39973.1; JOINED.  
DR EMBL; L21213; AAA39973.1; JOINED.  
DR EMBL; L21214; AAA39973.1; JOINED.  
DR EMBL; L21215; AAA39973.1; JOINED.  
DR EMBL; L21216; AAA39973.1; JOINED.  
DR EMBL; L21217; AAA39973.1; JOINED.  
DR EMBL; L21218; AAA39973.1; JOINED.  
DR EMBL; L21219; AAA39973.1; JOINED.  
DR EMBL; L21220; AAA39973.1; JOINED.  
DR EMBL; L21223; AAA39973.1; JOINED.  
DR HSSP; Q99405; IMPT.  
DR InterPro; IPR000209; Peptidase\_S8.  
DR InterPro; IPR002884; P\_domain.  
DR Pfam; PF01483; P; 1.  
DR Pfam; PF00082; Peptidase\_S8; 1.  
DR PRINTS; PR00723; SUBTILISIN.  
DR PRODOM; PD000717; P\_domain; 1.  
DR PROSITE; PS00136; SUBTILASE\_ASP; 1.  
DR PROSITE; PS00137; SUBTILASE\_HIS; 1.  
DR PROSITE; PS00138; SUBTILASE\_SER; 1.  
KW Serine protease  
SQ SEQUENCE 645 AA; 71980 MW; 54B07AA4A97D8AA0 CRC64;  
  
Query Match 52.1%; Score 37; DB 11; Length 645;  
Best Local Similarity 83.3%; Pred. No. 4.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 QQFWL 10  
Db 615 QQWWL 620  
|||:|  
RESULT 155  
Q9OU36 PRELIMINARY; PRT; 656 AA.  
AC Q9OU36;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
DE DNA, COMPLETE GENOME, ISOLATE:TLMV-CBD203.  
OS TTV-like mini virus.  
OC Viruses; ssDNA viruses; Circoviridae.  
OX NCBI\_TaxID=93678;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=TLMV-CBD203;  
RC Mishiro S.;  
RA Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RL

RN [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=TLMV-CBD203;  
RC Takahashi K., Iwasa Y., Hijikata M., Mishiro S.;  
RT "Identification of a new human DNA virus (TTV-like mini virus: TLMV)  
RT intermediately related to TT virus and chicken anemia virus.";  
RL Arch. Virol. 0:0-0(1999).  
DR EMBL; AB026929; BAA86945.1; -.  
SQ SEQUENCE 656 AA; 76983 MW; DA2CA1D3C2D83A37 CRC64;  
  
Query Match 52.1%; Score 37; DB 12; Length 656;  
Best Local Similarity 44.4%; Pred. No. 4.5e+02;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKQOWFW 9  
Db 268 KPQNNMFW 276  
|||:  
RESULT 156  
Q9PVX6 PRELIMINARY; PRT; 669 AA.  
AC Q9PVX6;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE CPEOMESODERMIN PROTEIN.  
GN CPEOMESODERMIN.  
OS Cynops pyrrhogaster (Japanese common newt).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Caudata; Salamandroides; Salamandridae; Cynops.  
OX NCBI\_TaxID=8330;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC TISSUE=EMBRYO;  
RX MEDLINE=99325981; PubMed=10400394;  
RA Sone K., Takahashi T.C., Takabatake Y., Takeshima K., Takabatake T.;  
RT "Expression of five novel T-box genes and brachyury during  
RT embryogenesis, and in developing and regenerating limbs and tails of  
RT newts.";  
RL Dev. Growth Differ. 41:321-333(1999).  
DR EMBL; AB019785; BAA84718.1; -.  
DR HSSP; P24781; 1XBR.  
DR InterPro; IPR001699; T-box.  
DR Pfam; PF00907; T-box; 1.  
DR PRINTS; PR00937; TBOX.  
DR SMART; SM00425; TBOX; 1.  
DR PROSITE; PS01283; TBOX\_1; 1.  
DR PROSITE; PS01264; TBOX\_2; 1.  
DR PROSITE; PS0252; TBOX\_3; 1.  
SQ SEQUENCE 669 AA; 72294 MW; 58D870CCF057CDEA CRC64;  
  
Query Match 52.1%; Score 37; DB 13; Length 669;  
Best Local Similarity 71.4%; Pred. No. 4.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 RPKQOWFW 8  
Db 489 PSPQWFW 495  
|||:  
RESULT 157  
Q9SRV5 PRELIMINARY; PRT; 765 AA.  
AC Q9SRV5;  
DT 01-MAY-2000 (TRENBLrel. 13, Created)  
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
DE PUTATIVE METHIONINE SYNTHASE.  
GN F20H23.19.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OL

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,  
RA Ronning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,  
RA Bowman C.B., White O., Nierman W.C., Fraser C.M.;  
RT "Arabidopsis thaliana chromosome III BAC F20H23 genomic sequence."  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC009540; AAF00639.1; -  
DR InterPro; IPR002629; Methionine\_synth.  
DR Pfam; PF01717; Methionine\_synth; 1  
SQ SEQUENCE 765 AA; 84583 MW; EB478D815910E701 CRC64;

Query Match 52.1%; Score 37; DB 10; Length 765;  
Best Local Similarity 66.7%; Pred. No. 5.2e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQWFW 9  
|||||  
DB 534 RPKPMTVFW 542

## RESULT 158

ID Q9LM03 PRELIMINARY; PRT; 765 AA.  
AC Q9LM03;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE METHIONINE SYNTHASE (EC 2.1.1.14).  
GN MS.  
OS Solanum tuberosum (potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zeh M., Leggewie G., Hoefgen R., Hesse H.;  
RT "Isolation and characterization of a cDNA encoding methionine synthase  
RT from Solanum tuberosum."  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF082893; AAF74393.1; -  
DR InterPro; IPR002629; Methionine\_synth.  
DR Pfam; PF01717; Methionine\_synth; 1.  
KW Transferase; Methyltransferase.  
SQ SEQUENCE 765 AA; 84665 MW; 6112AF7047DAD485 CRC64;

Query Match 52.1%; Score 37; DB 10; Length 765;  
Best Local Similarity 66.7%; Pred. No. 5.2e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQWFW 9  
|||||  
DB 534 RPKPMTVFW 542

## RESULT 159

ID Q9FFZ5 PRELIMINARY; PRT; 806 AA.  
AC Q9FFZ5;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE GB|AA|00669.1.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=COLUMBIA;  
RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,  
RA Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. XI."  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AP002544; BAB09687.1; -  
SQ SEQUENCE 806 AA; 89645 MW; 71645A5EE9E7746C CRC64;

Query Match 52.1%; Score 37; DB 10; Length 806;  
Best Local Similarity 50.0%; Pred. No. 5.5e+02;  
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQWFW 9  
|||  
DB 197 PEANEFW 204

## RESULT 160

Q9HEW7  
ID Q9HEW7 PRELIMINARY; PRT; 918 AA.  
AC Q9HEW7;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE NITROGEN RESPONSE FACTOR NRFL.  
OS Cladosporium fulvum (Fulvia fulva).  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;  
OC Dothideomycetes et Chaetothyriomycetes incertae sedis;  
OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.  
OX NCBI\_TaxID=5499;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Perez-Garcia A., Snoeijers S.S., Joosten M.H.A.J., Goosen T.,  
RA De Wit P.J.G.M.;  
RT "Expression of the avirulence gene Avr9 of the fungal tomato pathogen  
RT Cladosporium fulvum is regulated by the global nitrogen response  
RT factor NRFL."  
RL Mol. Plant Microbe Interact. 0:0-0(2001).  
DR EMBL; AF312694; AAG48616.1; -  
DR InterPro; IPR002965; P-rich\_extensn.  
DR InterPro; IPR000679; ZnF\_GATA.  
DR Pfam; PF00320; GATA; 1.  
DR PRINTS; PR00619; GATAZNFINGER.  
DR PRINTS; PR01217; PRICHEXTENS.  
DR SMART; SM00401; ZnF\_GATA; 1.  
DR PROSITE; PS00344; GATA\_ZN\_FINGER\_1; 1.  
DR PROSITE; PS01114; GATA\_ZN\_FINGER\_2; 1.  
SQ SEQUENCE 918 AA; 99077 MW; E98198D999BCA899 CRC64;

Query Match 52.1%; Score 37; DB 3; Length 918;  
Best Local Similarity 71.4%; Pred. No. 6.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQWFWL 10  
|||  
DB 908 PQEWEWL 914

## RESULT 161

O60043  
ID O60043 PRELIMINARY; PRT; 944 AA.  
AC O60043;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE NITROGEN RESPONSE REGULATOR.

```

GN NRRI.
OS Metarrhizium anisopliae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Clavicipitaceae; mitosporic Clavicipitaceae; Metarrhizium.
OX NCBI_TaxID=5530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MEL;
RA Screen S., Bailey A., Charnley K., Cooper R., Clarkson J.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ006468; CAA07052.1; -.
DR HSSP; P17429; 4GAT.
DR InterPro; IPR000679; ZNF_GATA.
DR Pfam; PF00320; GATA; 1.
DR PRINTS; PR00619; GATAZNFINGER.
DR SMART; SM00401; ZNF_GATA; 1.
DR PROSITE; PS00344; GATA_ZN_FINGER_1; 1.
DR PROSITE; PS00114; GATA_ZN_FINGER_2; 1.
SQ SEQUENCE 944 AA; 99686 MW; A1A723E658C23EA0 CRC64;

Query Match 52.1%; Score 37; DB 3; Length 944;
Best Local Similarity 71.4%; Pred. No. 6.4e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PQOWFWL 10
   ||| ||
Db 934 PQOWFWL 940

RESULT 162
Q9W570 PRELIMINARY; PRT; 1002 AA.
AC Q9W570;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2001 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2001 (TrEMBLrel. 17, Last annotation update)
DE DOR PROTEIN.
GN DOR OR EG-171E4.1 OR CG3093.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yeates D.L., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busan M.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matte B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

```

```

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL; AE003421; AAF45652.1; -.
DR FlyBase; FBgn0000482; dor.
DR InterPro; IPR000547; Clathrin_repeat.
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00299; CLH; 1.
DR SMART; SM00184; RING; 1.
SQ SEQUENCE 1002 AA; 115319 MW; 5B56BFE9040256BB CRC64;

Query Match 52.1%; Score 37; DB 5; Length 1002;
Best Local Similarity 71.4%; Pred. No. 6.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PQOWFWL 10
   ||| ||
Db 301 PQOWFWL 307

RESULT 163
Q9H8F3 PRELIMINARY; PRT; 1081 AA.
AC Q9H8F3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE CDNA FLJ13680 FIS, CLONE PLACE2000007, HIGHLY SIMILAR TO HOMO SAPIENS
DE KIA0913 PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isoqai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Naganari K., Masuno Y., Oshima A.;
RT "NEDO human cDNA sequencing project."
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK023742; BAB14664.1; -.
SQ SEQUENCE 1081 AA; 115370 MW; 01975A049C70A001 CRC64;

Query Match 52.1%; Score 37; DB 4; Length 1081;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 QNFWL 10
   |||||
Db 663 QNFWL 667

RESULT 164
Q94987 PRELIMINARY; PRT; 1301 AA.
AC Q94987;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

```

DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)  
DE KIAA0913 PROTEIN (FRAGMENT).  
GN KIAA0913.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
RN [1]  
NCBI\_TaxID=9606;  
SEQUENCE FROM N.A.  
RP TISSUE=BRAIN;  
RC MEDLINE=99156230; PubMed=10048485;  
RX Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirosewa M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 5:355-364(1998).  
DR EMBL; AB020720; BAA74936.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 1301 AA; 138677 MW; E0F0C4CE12615646 CRC64;  
Query Match 52.1%; Score 37; DB 4; Length 1301;  
Best Local Similarity 100.0%; Pred. No. 8.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 6 QWFWL 10  
Db 1134 QWFWL 1138  
RESULT 165  
O93457 PRELIMINARY; PRT; 1418 AA.  
AC O93457;  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE INSULIN-LIKE GROWTH FACTOR 1 RECEPTOR PRECURSOR.  
OS Scophthalmus maximus (turbot).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidae; Scophthalmidae; Scophthalmus.  
OX NCBI\_TaxID=52904;  
RN [1]  
SEQUENCE FROM N.A.  
RA Elies G., Duval H., Bonnet G., Wolff J., Boeuf G., Boujard D.;  
RT "Turbot insulin and insulin-like growth factor-1 receptors: cDNAs  
RT cloning and messenger RNAs expression during development.";  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE -> ADP + PROTEIN  
CC TYROSINE PHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (BY SIMILARITY).  
DR EMBL; AJ224993; CAA12278.1; -.  
DR HSP; P06213; IIRK.  
DR InterPro; IPR000494; EGFR\_L.  
DR InterPro; IPR000719; Euk.pkinase.  
DR InterPro; IPR003961; FN\_III.  
DR InterPro; IPR002174; Furin-like.  
DR InterPro; IPR002011; Receptor\_tyr\_kin\_II.  
DR InterPro; IPR001245; Tyr\_kin.  
DR Pfam; PF00041; fn3; 2.  
DR Pfam; PF00757; Furin-like; 1.  
DR Pfam; PF00069; pkinase; 1.  
DR Pfam; PF01030; Recep\_L\_domain; 1.  
DR PRINTS; PR00109; TYRKINASE.  
DR SMART; SM00060; FN3; 2.  
DR SMART; SM00219; TYRK; 1.  
DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.  
DR PROSITE; PS00109; PROTEIN\_KINASE\_TYR; 1.

DR PROSITE; PS00239; RECEPTOR\_TYR\_KIN\_II; 1.  
KW ATP-binding; Glycoprotein; Phosphorylation; Receptor; Signal;  
KW Transferase; Transmembrane; Tyrosine-protein kinase.  
FT SIGNAL 1 29 POTENTIAL.  
FT CHAIN 30 1418 INSULIN-LIKE GROWTH FACTOR 1 RECEPTOR.  
SQ SEQUENCE 1418 AA; 159826 MW; 5D9921332113C4AE CRC64;  
Query Match 52.1%; Score 37; DB 13; Length 1418;  
Best Local Similarity 85.7%; Pred. No. 9.5e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RPKPQQW 7  
Db 1099 RPKPQQW 1105  
RESULT 166  
Q91588 PRELIMINARY; PRT; 1589 AA.  
ID Q91588  
AC Q91588;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DE 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE COMPLEMENT COMPONENT C3 (FRAGMENT).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
SEQUENCE FROM N.A.  
RP TISSUE=LIVER;  
RC MEDLINE=95180322; PubMed=7875221;  
RA Lambiris J.D., Fappas J., Mavroidis M., Wang Y., Manzone H., Swager J.,  
RT Du Pasquier L., Silibovsky R.;  
RT "The third component of Xenopus complement: cDNA cloning, structural  
RT and functional analysis, and evidence for an alternate C3  
RT transcript.";  
RL Eur. J. Immunol. 25:572-578(1995).  
DR EMBL; U19253; AAB60608.1; -.  
DR HSP; P01024; IC3D.  
DR InterPro; IPR002890; A2M\_N.  
DR InterPro; IPR001599; Alpha\_2\_macrogllobln.  
DR InterPro; IPR000020; Anaphylatoxin.  
DR InterPro; IPR001840; Anaphylatoxn.  
DR InterPro; IPR001134; Netrin\_C.  
DR Pfam; PF00207; A2M; 1.  
DR Pfam; PF01835; A2M\_N; 1.  
DR Pfam; PF01821; ANATO; 1.  
DR Pfam; PF01759; NTR; 1.  
DR PRINTS; PR00004; ANAPHYLATOXN.  
DR SMART; SM00104; ANATO; 1.  
DR PROSITE; PS00477; ALPHA\_2\_MACROGLOBULIN; 1.  
DR PROSITE; PS01177; ANAPHYLATOXIN\_1; 1.  
DR PROSITE; PS01178; ANAPHYLATOXIN\_2; 1.  
FT NON\_TER 1 588 COMPLEMENT COMPONENT C3 BETA CHAIN.  
FT CHAIN 592 1489 COMPLEMENT COMPONENT C3 ALPHA CHAIN.  
FT CHAIN 1589 AA; 177904 MW; DCB777FB4B11456A CRC64;  
SQ SEQUENCE 1589 AA; 177904 MW; DCB777FB4B11456A CRC64;  
Query Match 52.1%; Score 37; DB 13; Length 1589;  
Best Local Similarity 66.7%; Pred. No. 1.1e+03;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 4 PQQWFW 9  
Db 691 PESWFW 696  
RESULT 167  
O17368

ID O17368 PRELIMINARY; PRT; 1635 AA.  
AC O17368;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE F48A11.1 PROTEIN.  
DE F48A11.1.  
GN Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=BRISTOL N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Boulton A.,  
Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,  
Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
Smaison N., Smith A., Sonhammer E., Staden K., Sulston J.,  
Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.,  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
elegans".  
RL Nature 368:32-38(1994).  
[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=BRISTOL N2;  
RA Bradshaw H.;  
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.  
[3]  
RN SEQUENCE FROM N.A.  
RP STRAIN=BRISTOL N2;  
RA Waterston R.;  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF026210; AAB71283.1; -  
SQ SEQUENCE 1635 AA; 186341 MW; F07C3281935A2E98 CRC64;  
  
Query Match 52.1%; Score 37; DB 5; Length 1635;  
Best Local Similarity 71.4%; Pred. No. 1.1e+03;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 4 PQQFWL 10  
Db 580 PQTWIL 586  
||| ||  
||| ||  
  
RESULT 168  
ID O51827 PRELIMINARY; PRT; 2458 AA.  
AC O51827;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE POLYKETIDE SYNTHASE TYPE I.  
GN PLTB.  
OS Pseudomonas fluorescens.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=294;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=PF-5;  
RX MEDLINE=98094250; PubMed=9434161;  
RA Nowak-Thompson B., Gould S.J., Loper J.E.;  
RT "Identification and sequence analysis of the genes encoding a  
polyketide synthase required for pyoluteorin biosynthesis in  
Pseudomonas fluorescens Pf-5.";  
RT Pseudomonas fluorescens Pf-5.";  
RL Gene 204:17-24(1997).

RN SEQUENCE FROM N.A.  
RP STRAIN=PF-5;  
RA Nowak-Thompson B., Gould S.J., Loper J.E.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF081920; AAC38075.1; -  
DR InterPro; IPR001227; Acyltransf\_domain.  
DR InterPro; IPR000794; Ketoacyl-synt.  
DR InterPro; IPR003880; Phosphopant\_attach.  
DR InterPro; IPR002155; Thiolase.  
DR Pfam; PF00698; Acyl\_transf; 1.  
DR Pfam; PF00109; ketoacyl-synt; 2.  
DR Pfam; PF00550; pp-binding; 2.  
DR PROSITE; PS00075; ACP\_DOMAIN; 2.  
DR PROSITE; PS00606; B\_KETOACYL\_SYNTHASE; 2.  
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN\_1.  
KW Phosphopantetheine; Transferase.  
SQ SEQUENCE 2458 AA; 262676 MW; AE756080AE1A5FB1 CRC64;  
  
Query Match 52.1%; Score 37; DB 2; Length 2458;  
Best Local Similarity 55.6%; Pred. No. 1.6e+03;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 RPKQOWFW 9  
Db 1613 RPAGQLWIW 1621  
||| |||  
||| |||  
  
RESULT 169  
ID Q99N19 PRELIMINARY; PRT; 523 AA.  
AC Q99N19;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE CYTOCHROME P450 CYP4F13.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=129/SV+P TYR C-CH TER/+; TISSUE=KIDNEY;  
RA Antonovic L., Kawashima H., Strobel H.;  
RT "Protein expression and catalytic activity assessment of mouse 4F  
clones.";  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF233643; AAK15009.1; -  
SQ SEQUENCE 523 AA; 59823 MW; 47E0E7840940CE21 CRC64;  
  
Query Match 51.4%; Score 36.5; DB 11; Length 523;  
Best Local Similarity 75.0%; Pred. No. 4.3e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
  
Qy 2 PKPQOWFW 9  
Db 55 PKP-SMFW 61  
||| |||  
||| |||  
  
RESULT 170  
ID Q99KY6 PRELIMINARY; PRT; 523 AA.  
AC Q99KY6;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE SIMILAR TO RIKEN CDNA I810054N16 GENE.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=MAMMARY TUMOR;

RA Strausberg R.;

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC003954; AAB03954.1; -

SQ SEQUENCE 523 AA; 59894 MW; 3927661E5FBD20CD CRC64;

Query Match

Best Local Similarity 51.4%; Score 36.5; DB 11; Length 523;

Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 PKPQOWFW 9

DB 55 PKP-SWFW 61

||| |||

RESULT 171

Q9SDN8

ID Q9SDN8 PRELIMINARY; PRT; 93 AA.

AC Q9SDN8;

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)

DE ELICITIN-LIKE PROTEIN (FRAGMENT).

OS Phytophthora capsici.

OC Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;

OC Phytophthora.

OX NCBI\_TaxID=4784;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=PI314-T1;

RA Bailey A.M.;

RL "Detection and identification of additional gene products induced by

RT the interaction between Phytophthora capsici and its host, Capsicum

RT annum.;"

RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF212434; AAF18483.1; -

FT NON\_TER 1

FT NON\_TER 93

SQ SEQUENCE 93 AA; 11050 MW; E57D17B614107060 CRC64;

Query Match

Best Local Similarity 50.7%; Score 36; DB 10; Length 93;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOW 7

DB 52 RPKVQOW 58

||| |||

RESULT 172

P73564

ID P73564 PRELIMINARY; PRT; 103 AA.

AC P73564;

DT 01-FEB-1997 (TREMBLrel. 02, Created)

DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)

DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)

DE HYPOTHETICAL 12.2 KDA PROTEIN.

GN SLR0881

OS Synechocystis sp. (strain PCC 6803).

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX NCBI\_TaxID=1148;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,

RA Miyajima N., Hirose M., Sugita M., Sugita M., Kimura T.,

RA Hosouchi T., Matsuno A., Muraki N., Nakazaki N., Naruo K., Okumura S.,

RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,

RA Tabata S.;

RT "Sequence analysis of the genome of the unicellular cyanobacterium

RT Synechocystis sp. strain PCC6803. II. Sequence determination of the

RT entire genome and assignment of potential protein-coding regions.;"

RL DNA Res 3:109-136(1996).

DR EMBL; D90907; BAAL7604.1; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 103 AA; 121174 MW; 24EE96F034055C71 CRC64;

Query Match

Best Local Similarity 50.7%; Score 36; DB 2; Length 103;

Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQWFW 9

DB 84 RPEHWYW 90

||: ||:

RESULT 173

O76548

ID O76548 PRELIMINARY; PRT; 104 AA.

AC O76548;

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)

DE GGTA (FRAGMENT).

GN GGTA.

OS Dictyostellium discoideum (Slime mold).

OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

OX NCBI\_TaxID=44689;

RN [1]

RP SEQUENCE FROM N.A.

RA Iranfar N., Loomis W.F.;

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF076603; AAC31542.1; -

FT NON\_TER 1

FT NON\_TER 1

SQ SEQUENCE 104 AA; 12263 MW; 6150084AE4714CFB CRC64;

Query Match

Best Local Similarity 50.7%; Score 36; DB 5; Length 104;

Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQWFW 9

DB 5 PQEWLW 10

||: ||:

RESULT 174

O9AFZ7

ID O9AFZ7 PRELIMINARY; PRT; 105 AA.

AC O9AFZ7;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE ORF, HYPOTHETICAL.

GN S0011.

OS Shigella flexneri.

OG Plasmid virulence pWR501.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Shigella.

OX NCBI\_TaxID=623;

RN [1]

RP SEQUENCE FROM N.A.

RA Venkatesan M.M., Goldberg M.B., Rose D.J., Grotbeck E.J., Burland V.,

RA Blattner F.R.;

RT "Complete DNA Sequence and Analysis of the Large Virulence Plasmid of

RT Shigella flexneri.;"

RL Infect. Immun. 0:0-0(2001).

DR EMBL; AF348706; AAK18322.1; -

KW Plasmid.

SQ SEQUENCE 105 AA; 11765 MW; 4D17F10F5FFBCBCC CRC64;

Query Match 50.7%; Score 36; DB 2; Length 105;  
 Best Local Similarity 60.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPQWFNL 10  
 |||||  
 Db 26 RPYDPQWFL 35

## RESULT 175

Q9DCS0 PRELIMINARY; PRT; 125 AA.  
 AC Q9DCS0  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE DETODINASE, IODOTHYRONINE, TYPE I.  
 GN DIOL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 ON NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=KIDNEY;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL; AK002549; BAB22180.1; -.  
 DR MGD; MGI:94896; Diol.  
 SQ SEQUENCE 125 AA; 14309 MW; 2C06DEC77C5F9E21 CRC64;

Query Match 50.7%; Score 36; DB 11; Length 125;  
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PQQWFL 10  
 |||||  
 Db 4 POLWLWL 10

## RESULT 176

Q36742 PRELIMINARY; PRT; 135 AA.  
 AC Q36742  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ATPASE 8 PROTEIN (FRAGMENT).  
 GN ATPASE 8.  
 OS Cottus kessleri.  
 OS Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Scorpaeniformes;  
 OC Cottoidel; Cottidae; Cottus.  
 ON NCBI\_TaxID=8099;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95287403; PubMed=7769616;  
 RA Slobodyanyuk S.Ja., Kirilchik S.V., Pavlova M.E., Belikov S.I.,  
 RA Novitsky A.L.;  
 RT "The evolutionary relationships of two families of cottoid fishes of  
 RT Lake Baikal (east Siberia) as suggested by analysis of mitochondrial  
 RT DNA."  
 RL J. Mol. Evol. 40:392-399(1995).  
 DR EMBL; S78299; AAD14277.1; -.  
 DR InterPro; IPR001421; ATP-synt\_8.  
 DR Pfam; PF00895; ATP-synt\_8; 1.  
 KW Mitochondrion.  
 FT NON\_TER 135 135  
 SQ SEQUENCE 135 AA; 15430 MW; B417E429E935357F CRC64;

Query Match 50.7%; Score 36; DB 8; Length 135;  
 Best Local Similarity 45.5%; Pred. No. 1.4e+02;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RPKPQWFNL 11  
 :||:|  
 Db 44 RPKTEPTWPM 54

## RESULT 177

Q9RVC1 PRELIMINARY; PRT; 154 AA.  
 ID Q9RVC1  
 AC Q9RVC1  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL 16.7 KDA PROTEIN.  
 GN DR1108.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Thermococcus group; Deinococcales; Deinococcus.  
 ON NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RL.  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
 RA Vanathavan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.;  
 RT "Genome sequence of the radioresistant bacterium Deinococcus  
 RT radiodurans RL."  
 RL Science 286:1571-1577(1999).  
 DR EMBL; AE001961; AAF10689.1; -.  
 DR TIGR; DR1108; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 154 AA; 16714 MW; D888E4ACFFE70234 CRC64;

Query Match 50.7%; Score 36; DB 2; Length 154;  
 Best Local Similarity 54.5%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPQWFNL 11  
 |:|||  
 Db 46 ROPQTAFWLL 56

## RESULT 178

```
Q9HWX8          PRELIMINARY;      PRT;    171 AA.
AC Q9HWX8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE PHOSPHATIDYLGLYCEROPHOSPHATASE A.
GN PG OR PA4050.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PA01.
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Huynh W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE004821; AAG07437.1; -.
KW Complete proteome.
SQ SEQUENCE 171 AA; 19606 MW; 4C8D5C86276F892D CRC64;

Query Match      50.7%; Score 36; DB 2; Length 171;
Best Local Similarity 50.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQWFWM 11
DB 107 PEGWWLL 114

RESULT 179
Q9FRE0          PRELIMINARY;      PRT;    241 AA.
AC Q9FRE0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE HYPOTHETICAL 25.1 KDA PROTEIN.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,
RA Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
RA Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNBa0013M12 genomic sequence."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC082644; AA046138.1; -.
KW Hypothetical protein.
SQ SEQUENCE 241 AA; 25126 MW; F07B4F4BF05DB6AC CRC64;
```

```
Query Match      50.7%; Score 36; DB 10; Length 241;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QQWFWM 10
DB 149 QQWLWL 154
```

```
RESULT 180
Q9TAK3          PRELIMINARY;      PRT;    250 AA.
AC Q9TAK3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE NADH DEHYDROGENASE SUBUNIT 6 (EC 1.6.5.3).
GN NAD6.
OS Cafeteria roenbergensis.
OC Mitochondrion.
OC Eukaryota; stramenopiles; Bicosoecida; Cafeteria.
OX NCBI_TaxID=33653;
RN [1]
RP SEQUENCE FROM N.A.
RA Burger G.;
RT "The mitochondrial genome of Cafeteria roenbergensis."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF193903; AAF05783.1; -.
DR InterPro; IPR001457; Oxidored_q3.
DR Pfam; PF00459; Oxidored_q3; 1.
KW Oxidoreductase; Mitochondrion.
SQ SEQUENCE 250 AA; 29049 MW; E992796F89D8255B CRC64;

Query Match      50.7%; Score 36; DB 8; Length 250;
Best Local Similarity 62.5%; Pred. No. 2.5e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQQWF 9
DB 235 PKPKGFFW 242
```

```
RESULT 181
O19110          PRELIMINARY;      PRT;    254 AA.
AC O19110;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE TESTIS-SPECIFIC PROTEIN (FRAGMENT).
GN TSPY.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTES;
RX MEDLINE=97341111; PubMed=9195993;
RA Vogel T., Dechend F., Manz E., Jung C., Jakubiczka S., Fehr S.,
RA Schmidtke J., Schmieders F.;
RT "Organization and expression of bovine TSPY."
RL Mamm. Genome 8:491-496(1997).
DR EMBL; U75896; AAB72143.1; -.
DR InterPro; IPR002164; NAP_family.
DR Pfam; PF00956; NAP_family; 1.
FT NON_TER 1
SQ SEQUENCE 254 AA; 29745 MW; B36A29399A9132DC CRC64;
```

```
Query Match      50.7%; Score 36; DB 6; Length 254;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9
DB 186 RSTPVHFW 194
```



```

RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Ragiol I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT *DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae";
RL Nature 406:477-483(2000).
DR EMBL: AE004233; AAF94705.1; -.
DR TIGR: VC1551; -.
DR InterPro: IPR000515; BPD_transp.
DR Pfam: PF00528; BPD_transp; 1.
KW Complete proteome.
SQ SEQUENCE 280 AA; 31666 MW; C12BF840FC2742CC CRC64;

Query Match 50.7%; Score 36; DB 2; Length 280;
Best Local Similarity 50.0%; Pred. NO. 2.8e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PKPQQWFWM 11
DB 105 PYASAWFWLI 114
| | | | |
DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR0500.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.J., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT *Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.;
RL Science 286:1571-1577(1999).
DR EMBL: AE001909; AAF10080.1; -.
DR TIGR: DR0500; -.
DR InterPro: IPR003797; DUF194.
DR Pfam: PF02645; DUF194; 1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 30591 MW; 4E247F9B2D278771 CRC64;

Query Match 50.7%; Score 36; DB 2; Length 287;
Best Local Similarity 71.4%; Pred. NO. 2.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQW 7
DB 64 QPSPPQW 70
| | | | |
DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR0500.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.J., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT *Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.;
RL Science 286:1571-1577(1999).
DR EMBL: AE001909; AAF10080.1; -.
DR TIGR: DR0500; -.
DR InterPro: IPR003797; DUF194.
DR Pfam: PF02645; DUF194; 1.
KW Complete proteome.
SQ SEQUENCE 287 AA; 30591 MW; 4E247F9B2D278771 CRC64;

Query Match 50.7%; Score 36; DB 5; Length 268;
Best Local Similarity 55.6%; Pred. NO. 2.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9
DB 116 RHRPDQGW 124
| | | | |
DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR0500.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.J., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT *Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.;
RL Science 286:1571-1577(1999).
DR EMBL: AE001909; AAF10080.1; -.
DR TIGR: DR0500; -.
DR InterPro: IPR003797; DUF194.
DR Pfam: PF02645; DUF194; 1.
KW Complete proteome.
SQ SEQUENCE 268 AA; 30361 MW; 3AE5FCC8E84F6783 CRC64;

Query Match 50.7%; Score 36; DB 5; Length 268;
Best Local Similarity 55.6%; Pred. NO. 2.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9
DB 116 RHRPDQGW 124
| | | | |
DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR0500.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.J., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT *Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.;
RL Science 286:1571-1577(1999).
DR EMBL: AE001909; AAF10080.1; -.
DR TIGR: DR0500; -.
DR InterPro: IPR003797; DUF194.
DR Pfam: PF02645; DUF194; 1.
KW Complete proteome.
SQ SEQUENCE 268 AA; 30361 MW; 3AE5FCC8E84F6783 CRC64;

Query Match 50.7%; Score 36; DB 5; Length 268;
Best Local Similarity 55.6%; Pred. NO. 2.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9
DB 116 RHRPDQGW 124
| | | | |
DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR0500.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.J., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT *Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.;
RL Science 286:1571-1577(1999).
DR EMBL: AE001909; AAF10080.1; -.
DR TIGR: DR0500; -.
DR InterPro: IPR003797; DUF194.
DR Pfam: PF02645; DUF194; 1.
KW Complete proteome.
SQ SEQUENCE 268 AA; 30361 MW; 3AE5FCC8E84F6783 CRC64;

Query Match 50.7%; Score 36; DB 5; Length 268;
Best Local Similarity 55.6%; Pred. NO. 2.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9
DB 116 RHRPDQGW 124
| | | | |
DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR0500.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.J., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT *Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.;
RL Science 286:1571-1577(1999).
DR EMBL: AE001909; AAF10080.1; -.
DR TIGR: DR0500; -.
DR InterPro: IPR003797; DUF194.
DR Pfam: PF02645; DUF194; 1.
KW Complete proteome.
SQ SEQUENCE 268 AA; 30361 MW; 3AE5FCC8E84F6783 CRC64;

Query Match 50.7%; Score 36; DB 5; Length 268;
Best Local Similarity 55.6%; Pred. NO. 2.7e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9
DB 116 RHRPDQGW 124
| | | | |
DE CONSERVED HYPOTHETICAL PROTEIN.
GN DR0500.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.J., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT *Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.;
RL Science 286:1571-1577(1999).
DR EMBL: AE001909; AAF10080.1; -.
DR TIGR: DR0500; -.
```

```

ID Q53945 PRELIMINARY; PRT; 302 AA.
AC Q53945;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE BACTERIOPHAGE (PHIC31) RESISTANCE (PGLY AND PGL2) GENES, COMPLETE
DE CDS'5 (FRAGMENT).
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RX MEDLINE=95370146; PubMed=7642495;
RA Bedford D.J., Laity C., Buttner M.J.;
RT "Two genes involved in the phase-variable phi C31 resistance mechanism
RT of Streptomyces coelicolor A3(2).";
RL J. Bacteriol. 177:4681-4689(1995).
RL EMBL; L37531; AAB00368.1; -.
DR InterPro; IPR001969; Asp.protease.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 302 AA; 34027 MW; 7E9DFC5A218402D7 CRC64;

Query Match 50.7%; Score 36; DB 2; Length 302;
Best Local Similarity 57.1%; Pred. No. 3e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQWF 9
I :|||
DB 93 KTEQWY 99

RESULT 186
Q9CYX6 PRELIMINARY; PRT; 320 AA.
ID Q9CYX6;
AC Q9CYX6;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE WINGLESS-RELATED MTW INTEGRATION SITE 5B.
GN WNT5B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stauble F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALLING

```

```

CC MOLECULE WHICH AFFECT THE DEVELOPMENT OF DISCRETE REGIONS OF
CC TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE
CC -!- EXTRACELLULAR MATRIX (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHER MEMBERS OF THE WNT FAMILY.
DR EMBL; AK013218; BAB28720.1; -.
DR MGD; MGI:98959; Wnt5b.
DR InterPro; IPR000970; Wnt1.
DR Pfam; PF00110; wnt; 1.
DR PRINTS; PRQ1349; WNTPROTEIN.
DR SMART; SM00097; WNT1; 1.
DR PROSITE; PS00246; WNT1; 1.
KW Developmental protein; Glycoprotein.
SQ SEQUENCE 320 AA; 35600 MW; DF75DEDF755B139B CRC64;

Query Match 50.7%; Score 36; DB 11; Length 320;
Best Local Similarity 54.5%; Pred. No. 3.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 RPK--POQWF 9
||| :|||
DB 161 RPKDLPDRLW 171

RESULT 187
Q9B4H0 PRELIMINARY; PRT; 343 AA.
ID Q9B4H0;
AC Q9B4H0;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE CYTOCHROME B (FRAGMENT).
OS Callisaurus draconoides.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Phrynosomatinae;
OC Callisaurus.
OX NCBI_TaxID=43586;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ROM4146;
RX MEDLINE=21175753; PubMed=11277627;
RA Trepanier T.L., Murphy R.W.;
RT "The Coachella Valley Fringe-Toed Lizard (Uma inornata): Genetic
RT Diversity and Phylogenetic Relationships of an Endangered Species.";
RL Mol. Phylogenet. Evol. 18:327-334(2001).
DR EMBL; AF302008; AAK32108.1; -.
KW Mitochondrion.
FT NON_TER 1
FT NON_TER 343
SQ SEQUENCE 343 AA; 38734 MW; 812A8815D6A2E331 CRC64;

Query Match 50.7%; Score 36; DB 8; Length 343;
Best Local Similarity 54.5%; Pred. No. 3.4e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQWFWM 11
||| :|||
DB 282 RPKSOTMFWLL 292

RESULT 188
Q47471 PRELIMINARY; PRT; 347 AA.
ID Q47471;
AC Q47471;
DT 01-NOV-1996 (TEMBLrel. 01, Created)
DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)
DT 01-JUN-2000 (TEMBLrel. 14, Last annotation update)
DE PECTATE LIASE.
GN PELB.

```

```

OS Erwinia carotovora.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pectobacterium.
OX NCBI_TaxID=554;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SCC3193;
RA Heikinheimo R., Flego D., Pirhonen M., Karlsson M.B., Eriksson A.,
RA Mae A., Koiv V., Paiva E.T.;
RT "Characterization of a novel pectate lyase from Erwinia carotovora
RT subsp. carotovora.";
RL Mol. Plant Microbe Interact. 8:207-217(1995).
DR EMBL; X79232; CAA55814.1; -.
KW Lyase.
SQ SEQUENCE 347 AA; 37432 MW; 3E70EECB120D799. CRC64;

Query Match 50.7%; Score 36; DB 2; Length 347;
Best Local Similarity 44.4%; Pred. No. 3.5e+02;
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 2 PKPQOWFWL 10
Db 82 PKSDYWWV 90

RESULT 189
ID Q9MM76 PRELIMINARY; PRT; 347 AA.
AC Q9MM76;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CYTOCHROME B (FRAGMENT).
GN CYTB.
OS Gallotia galloti.
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Scincomorpha; Lacertoidea;
OC Lacertidae; Gallotia.
OX NCBI_TaxID=39310;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu J.;
RT "Toward the phylogeny of the family Lacertidae - why 4708 base pairs
RT of mtDNA sequences cannot draw the picture.";
RL Biol. J. Linn. Soc. Lond. 71:203-217(2000).
CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE
CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A
CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL
CC COUPLED TO ATP SYNTHESIS (BY SIMILARITY).
CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY
CC BOUND TO THE PROTEIN (BY SIMILARITY).
CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.
DR EMBL; AF206534; AAF70424.1; -.
DR InterPro; IPR001179; Cyt_b.b6.
DR Pfam; PF00033; cytochrome_b.c1.
DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
DR PROSITE; PS00193; CYTOCHROME_B_QQ; UNKNOWN_1.
KW Electron transport; Heme; Mitochondrion; Respiratory chain;
KW Transmembrane.
FT NON_TER 1
SQ SEQUENCE 347 AA; 39168 MW; 258F788D4485A139. CRC64;

Query Match 50.7%; Score 36; DB 8; Length 347;
Best Local Similarity 54.5%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

Oy 1 RPKPOOWFWLM 11
Db 286 RPKSOMLFWLL 296

RESULT 190
ID Q9BV04 PRELIMINARY; PRT; 359 AA.
AC Q9BV04;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SIMILAR TO WINGLESS-RELATED MMTV INTEGRATION SITE 5B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=RHADOMYOSARCOMA;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC001749; AA01749.1; -.
SQ SEQUENCE 359 AA; 40323 MW; 6E35EE2B0AF1FD29. CRC64;

Query Match 50.7%; Score 36; DB 4; Length 359;
Best Local Similarity 54.5%; Pred. No. 3.6e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Oy 1 RPK--POOWFW 9
Db 148 RPKDLPRDMLW 158

RESULT 191
ID Q9XT34 PRELIMINARY; PRT; 360 AA.
AC Q9XT34;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ALPHA-1,3-FUCOSYLTRANSFERASE (FRAGMENT).
GN FUT3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Meijerink E., Voegel P., Stranzinger G.;
RT "Sus scrofa alpha(1,3)fucosyltransferase gene.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF130972; AAD39143.1; -.
DR InterPro; IPR001503; Glyco.transf_10.
DR Pfam; PF00852; Glyco.transf_10; 1.
KW Transferase; Glycosyltransferase.
FT NON_TER 1
SQ SEQUENCE 360 AA; 42160 MW; 46405E5AF9EE7A3A. CRC64;

Query Match 50.7%; Score 36; DB 6; Length 360;
Best Local Similarity 55.6%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 RPKPOOWFW 9
Db 125 RPPGQWVW 133

RESULT 192
ID O35886 PRELIMINARY; PRT; 362 AA.

```

AC Q35886;  
DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE ALPHA(1.3)FUCOSYLTRANSFERASE.  
GN ALPHA.  
OS Cricetus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetus.  
OX NCBI\_TaxID=10029;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang A., Potvin B., Kumar R., Zaiman A., Stanley P.;  
RL MEDLINE-99214609; PubMed-10187834;  
DR EMBL: U78737; AAB64355.1; -.  
DR InterPro: IPR001503; Glyco\_transf\_10.  
DR Pfam: PF00852; Glyco\_transf\_10; 1.  
KW Transferase; Glycosyltransferase.  
SQ SEQUENCE 362 AA; 41810 MW; A67940D57D47004C CRC64;  
  
Query Match 50.7%; Score 36; DB 11; Length 362;  
Best Local Similarity 55.6%; Pred. No. 3.6e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKPQQWF 9  
|||  
DB 127 RPPGQRWV 135  
  
RESULT 193  
Q9R220  
ID Q9R220 PRELIMINARY; PRT; 362 AA.  
AC Q9R220;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE ALPHA(1.3)FUCOSYLTRANSFERASE 6A.  
GN FUT6A.  
OS Cricetus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetus.  
OX NCBI\_TaxID=10029;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang A., Potvin B., Zaiman A., Chen W., Kumar R., Phillips L.,  
RA Stanley P.;  
RT "The gain-of-function Chinese hamster ovary mutant LEC1B expresses  
one of two Chinese hamster FUT6 genes due to the loss of a negative  
regulatory factor.";  
RL J. Biol. Chem. 274:10439-10450(1999).  
DR EMBL: AF090450; AAD24888.1; -.  
DR InterPro: IPR001503; Glyco\_transf\_10.  
DR Pfam: PF00852; Glyco\_transf\_10; 1.  
KW Transferase; Glycosyltransferase.  
SQ SEQUENCE 362 AA; 41767 MW; 2256EA145B03DA13 CRC64;  
  
Query Match 50.7%; Score 36; DB 11; Length 362;  
Best Local Similarity 55.6%; Pred. No. 3.6e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKPQQWF 9  
|||  
DB 127 RPPGQRWV 135  
  
RESULT 194  
Q9R219  
ID Q9R219 PRELIMINARY; PRT; 362 AA.

AC Q9R219;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE ALPHA(1.3)FUCOSYLTRANSFERASE 6B.  
GN FUT6B.  
OS Cricetus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetus.  
OX NCBI\_TaxID=10029;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang A., Potvin B., Zaiman A., Chen W., Kumar R., Phillips L.,  
RA Stanley P.;  
RT "The gain-of-function Chinese hamster ovary mutant LEC1B expresses  
one of two Chinese hamster FUT6 genes due to the loss of a negative  
regulatory factor.";  
RL J. Biol. Chem. 274:10439-10450(1999).  
DR EMBL: AF090449; AAD24887.1; -.  
DR InterPro: IPR001503; Glyco\_transf\_10.  
DR Pfam: PF00852; Glyco\_transf\_10; 1.  
KW Transferase; Glycosyltransferase.  
SQ SEQUENCE 362 AA; 41743 MW; 480D106C40DE5F30 CRC64;  
  
Query Match 50.7%; Score 36; DB 11; Length 362;  
Best Local Similarity 55.6%; Pred. No. 3.6e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKPQQWF 9  
|||  
DB 127 RPPGQRWV 135  
  
RESULT 195  
P73843  
ID P73843 PRELIMINARY; PRT; 369 AA.  
AC P73843;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE HYPOTHETICAL 42.2 KDA PROTEIN.  
GN SLL1611.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97061201; PubMed-8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimpo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
Synechocystis sp. strain PCC6803. II. Sequence determination of the  
entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL: D90910; BAA17900.1; -.  
DR InterPro: IPR001225; FA\_desaturase.  
DR ProDom: PD001081; FA\_desaturase; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 369 AA; 42180 MW; 25F3FDE2250520F9 CRC64;  
  
Query Match 50.7%; Score 36; DB 2; Length 369;  
Best Local Similarity 71.4%; Pred. No. 3.7e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 4 PQQWFL 10  
|||

Db 52 PGQWLWL 58

RESULT 196

Q9XPE2 PRELIMINARY; PRT; 380 AA.

AC Q9XPE2; 12, Created)

DT 01-NOV-1999 (TREMELREL. 12, Last sequence update)

DT 01-NOV-1999 (TREMELREL. 12, Last sequence update)

DE 01-JUN-2001 (TREMELREL. 17, Last annotation update)

DE CYTOCHROME B.

GN CYTB.

OS Eumeces egregius.

OG Mitochondrion.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Lepidosauria; Squamata; Scieroglossa; Scincomorpha; Scincoidae;

OC Scincidae; Eumeces.

OX NCBI\_TaxID=52436;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99297345; PubMed=10368956;

RA Kumazawa Y., Nishida M.;

RT "Complete mitochondrial DNA sequences of the green turtle and blue-tailed mole skink: statistical evidence for archosaurian affinity of turtles.";

RT Mol. Biol. Evol. 16:784-792(1999).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=96073446; PubMed=7476123;

RA Kumazawa Y., Nishida M.;

RT "Variations in mitochondrial trna gene organization of reptiles as phylogenetic markers.";

RL Mol. Biol. Evol. 12:759-772(1995).

CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL COUPLED TO ATP SYNTHESIS (BY SIMILARITY).

CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY BOUND TO THE PROTEIN (BY SIMILARITY).

CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B, CYTOCHROME C1 AND THE RIESKE PROTEIN (BY SIMILARITY).

CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B/B6 FAMILY.

DR EMBL; AB016606; BAA79222.1; -.

DR InterPro: IPR000179; Cyt\_b.b6.

DR Pfam: PF00032; cytochrome\_b\_c; 1.

DR Pfam: PF00033; cytochrome\_b\_n; 1.

DR PROSITE; PS00192; CYTOCHROME\_B\_HEME; 1.

DR PROSITE; PS00193; CYTOCHROME\_B\_QO; UNKNOWN1.

KW Electron transport; Heme; Mitochondrion; Respiratory chain; Transmembrane.

SQ SEQUENCE 380 AA; 42527 MW; 81844197B9CFB471 CRC64;

Query Match 50.7%; Score 36; DB 8; Length 380;

Best Local Similarity 54.5%; Pred. No. 3.8e+02;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQQWFILM 11

Db 319 RPASQTMFWLL 329

RESULT 197

Q9YGX6 PRELIMINARY; PRT; 385 AA.

AC Q9YGX6;

DT 01-MAY-1999 (TREMELREL. 10, Created)

DT 01-MAY-1999 (TREMELREL. 10, Last sequence update)

DE 01-JUN-2001 (TREMELREL. 17, Last annotation update)

DE WNT-5A.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

Gallus.

OX NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99372672; PubMed=10445500;

RA Kawakami Y., Wada N., Nishimatsu S., Ishikawa T., Noji S., Nohno T.;

RT "Involvement of Wnt-5a in chondrogenic pattern formation in the chick limb bud.";

RL Dev. Growth Differ. 41:29-40(1999).

CC -!- FUNCTION: PROBABLE DEVELOPMENTAL PROTEIN. MAY BE A SIGNALLING MOLECULE WHICH AFFECT THE DEVELOPMENT OF DISCRETE REGIONS OF TISSUES. IS LIKELY TO SIGNAL OVER ONLY FEW CELL DIAMETERS (BY SIMILARITY).

CC -!- SUBCELLULAR LOCATION: POSSIBLY SECRETED AND ASSOCIATES WITH THE EXTRACELLULAR MATRIX.

CC -!- SIMILARITY: TO OTHER MEMBERS OF THE WNT FAMILY.

DR EMBL; AB006014; BAA75242.1; -.

DR InterPro: IPR000970; Wnt1.

DR Pfam: PF00110; wnt; 1.

DR PRINTS; PR01349; WNTPROTEIN.

DR SMART; SM00097; WNT1; 1.

DR PROSITE; PS00246; WNT1; 1.

KW Developmental protein; Glycoprotein.

SQ SEQUENCE 385 AA; 43005 MW; 409F7440368B2360 CRC64;

Query Match 50.7%; Score 36; DB 13; Length 385;

Best Local Similarity 54.5%; Pred. No. 3.8e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 RPK--PQQWF 9

Db 174 RPKDLPRDLW 184

RESULT 198

Q99JJ4 PRELIMINARY; PRT; 418 AA.

AC Q99JJ4;

DT 01-JUN-2001 (TREMELREL. 17, Created)

DT 01-JUN-2001 (TREMELREL. 17, Last sequence update)

DT 01-JUN-2001 (TREMELREL. 17, Last annotation update)

DE SIMILAR TO KIAA0317 GENE PRODUCT (FRAGMENT).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC006074; AAB06074.1; -.

FT NON\_TER 1

SQ SEQUENCE 418 AA; 47675 MW; B95D5FF0D0AE6863 CRC64;

Query Match 50.7%; Score 36; DB 11; Length 418;

Best Local Similarity 55.6%; Pred. No. 4.1e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9

Db 324 REKVMRWF 332

RESULT 199

Q9VNP0 PRELIMINARY; PRT; 454 AA.

AC Q9VNP0;

DT 01-MAY-2000 (TREMELREL. 13, Created)

DT 01-MAY-2000 (TREMELREL. 13, Last sequence update)

DT 01-MAY-2000 (TREMELREL. 13, Last annotation update)

DE CG1169 PROTEIN.  
 GN CG1169.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 ON NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.C., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector K., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003600; AAF51889.1; -  
 DR FlyBase; FBgn0037428; CG1169.  
 SQ SEQUENCE 454 AA; 51320 MW; A75AAAD97E716573 CRC64;

Query Match 50.7%; Score 36; DB 5; Length 454;  
 Best Local Similarity 75.0%; Pred. NO. 4.5e+02;  
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKQQWF 8  
 ||| |||  
 Db 53 RPKALQWF 60

RESULT 200  
 Q9MGA9  
 ID Q9MGA9 PRELIMINARY; PRT; 482 AA.  
 AC Q9MGA9;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DE NADH DEHYDROGENASE SUBUNIT 4 (EC 1.6.5.3).  
 GN NAD4.  
 OS Chrysodidymus synuroideus.

OG Mitochondrion.  
 OC Eukaryota; Stramenopiles; Chrysophyceae; Synurales; Chrysodidymus.  
 ON NCBI\_TaxID=47573;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Chesnick J.M., Goff M., Graham J., Ocampo C., Lang B.F., Self E.,  
 RA Burger G.;  
 RT "The mitochondrial genome of the stramenopile alga, Chrysodidymus  
 RT synuroideus. Complete sequence, gene content and genome  
 RT organization."  
 RL Nucleic Acids Res. 0:0-0(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Burger G.;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
 CC -1- SIMILARITY: TO NADH-UBIQUINONE/PLASTOQUINONE (COMPLEX I), VARIOUS  
 CC CHAINS.  
 DR EMBL; AF222718; AAF36940.1; -  
 DR InterPro; IPR001750; Oxidored\_q1.  
 DR Pfam; PF00361; oxidored\_q1; 1.  
 KW Mitochondrion; NAD; Oxidoreductase; Ubiquinone.  
 SQ SEQUENCE 482 AA; 55109 MW; 731DFC959E4A7574 CRC64;

Query Match 50.7%; Score 36; DB 8; Length 482;  
 Best Local Similarity 83.3%; Pred. NO. 4.7e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 QQFWL 10  
 ||| ||  
 Db 203 QQWLWL 208

Search completed: April 1, 2002, 16:20:10  
 Job time: 156 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2002, 16:17:34 ; Search time 13.22 Seconds

(without alignments)  
30.508 Million cell updates/sec

Title: US-09-988-792-1

Perfect score: 61

Sequence: 1 RPKPQQFFGLM 11

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 127

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 50%

Maximum Match 100%

Listing first 1000 summaries

Database :

SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID      | Description         |
|------------|-------|---------------|--------|------------|---------------------|
| 1          | 61    | 100.0         | 11     | TKNA_HORSE | P01290 equus caball |
| 2          | 61    | 100.0         | 115    | TKN1_RABIT | P41540 oryctolagus  |
| 3          | 61    | 100.0         | 129    | TKN1_HUMAN | P20366 homo sapien  |
| 4          | 61    | 100.0         | 130    | TKN1_BOVIN | P01289 bos taurus   |
| 5          | 61    | 100.0         | 130    | TKN1_MESAU | Q60541 mesocricetu  |
| 6          | 61    | 100.0         | 130    | TKN1_MOUSE | P41539 mus musculus |
| 7          | 61    | 100.0         | 130    | TKN1_RAT   | P06767 rattus norv  |
| 8          | 58    | 95.1          | 11     | TKNA_CHICK | P19850 gallus gall  |
| 9          | 50    | 82.0          | 11     | TKNA_ONCMY | P28499 oncorhynch   |
| 10         | 49    | 80.3          | 11     | TKNA_GADMO | P28498 gadus morhu  |
| 11         | 48    | 78.7          | 11     | TKNA_SCYCA | P41333 scyllorhinu  |
| 12         | 44    | 72.1          | 11     | TKN4_PSEGU | P42989 pseudophryn  |
| 13         | 44    | 72.1          | 11     | TKN5_PSEGU | P42990 pseudophryn  |
| 14         | 44    | 72.1          | 12     | TKN1_KASMA | P08613 kassina mac  |
| 15         | 43    | 70.5          | 11     | TKNA_RANRI | P29207 rana ridibu  |
| 16         | 41    | 67.2          | 11     | TKNA_RANCA | P22688 rana catesb  |
| 17         | 41    | 67.2          | 3828   | TRX_DROVI  | Q24742 drosophila   |
| 18         | 37    | 60.7          | 11     | TKN1_UPERU | P08612 uperoleia r  |
| 19         | 37    | 60.7          | 11     | TKN2_PSEGU | P42987 pseudophryn  |
| 20         | 36    | 59.0          | 11     | TKN1_PSEGU | P42986 pseudophryn  |
| 21         | 36    | 59.0          | 11     | TKN3_PSEGU | P42988 pseudophryn  |
| 22         | 36    | 59.0          | 12     | TKN2_KASMA | P08614 kassina mac  |
| 23         | 36    | 59.0          | 12     | TKN_KASSE  | P08611 kassina sen  |
| 24         | 36    | 59.0          | 167    | SERO_GALME | O76192 gallieria me |
| 25         | 36    | 59.0          | 313    | ISPE_HAEIN | P45271 haemophilus  |
| 26         | 36    | 59.0          | 728    | EF2_ARCFU  | O28385 archaeoglob  |
| 27         | 35    | 57.4          | 512    | SYK_VIBCH  | Q8ku60 vibrio chol  |
| 28         | 35    | 57.4          | 518    | CP3R_ONCMY | O42563 oncorhynch   |
| 29         | 35    | 57.4          | 1092   | RELA_MYXXA | O52177 myxococcu    |
| 30         | 35    | 57.4          | 1093   | DP2L_METTH | O27579 methanobact  |
| 31         | 34    | 55.7          | 493    | Y130_MYCPN | P75506 mycoplasma   |
| 32         | 34    | 55.7          | 498    | C6B1_PAPPO | Q04552 papilio pol  |
| 33         | 34    | 55.7          | 503    | CP3A_MESAU | Q64148 mesocricetu  |

|            |   |      |      |    |    |            |
|------------|---|------|------|----|----|------------|
| CP31_RAT   | 1 | 504  | 55.7 | 34 | 34 | CP31_RAT   |
| PPOE_LYCES | 1 | 585  | 55.7 | 34 | 34 | PPOE_LYCES |
| PPOE_LYCES | 1 | 587  | 55.7 | 34 | 34 | PPOE_LYCES |
| PPOB_SOLFU | 1 | 588  | 55.7 | 34 | 34 | PPOB_SOLFU |
| XJB0_YEAST | 1 | 666  | 55.7 | 34 | 34 | XJB0_YEAST |
| Y025_CABEL | 1 | 1799 | 55.7 | 34 | 34 | Y025_CABEL |
| YK00_CABEL | 1 | 282  | 54.1 | 33 | 33 | YK00_CABEL |
| CAG5_CHICK | 1 | 404  | 54.1 | 33 | 33 | CAG5_CHICK |
| CRFL_CHICK | 1 | 420  | 54.1 | 33 | 33 | CRFL_CHICK |
| SYK_STAAG  | 1 | 495  | 54.1 | 33 | 33 | SYK_STAAG  |
| C6B3_PAPPO | 1 | 498  | 54.1 | 33 | 33 | C6B3_PAPPO |
| SYK_PASMU  | 1 | 501  | 54.1 | 33 | 33 | SYK_PASMU  |
| SYK_HAEIN  | 1 | 502  | 54.1 | 33 | 33 | SYK_HAEIN  |
| SVK1_ECOLI | 1 | 504  | 54.1 | 33 | 33 | SVK1_ECOLI |
| SVK2_ECOLI | 1 | 504  | 54.1 | 33 | 33 | SVK2_ECOLI |
| FRK_HUMAN  | 1 | 505  | 54.1 | 33 | 33 | FRK_HUMAN  |
| SYK_ACICA  | 1 | 509  | 54.1 | 33 | 33 | SYK_ACICA  |
| COX1_APILI | 1 | 521  | 54.1 | 33 | 33 | COX1_APILI |
| IMAI_ARATH | 1 | 596  | 54.1 | 33 | 33 | IMAI_ARATH |
| TLD_DROME  | 1 | 1057 | 54.1 | 33 | 33 | TLD_DROME  |
| ODPB_BACST | 1 | 324  | 53.3 | 34 | 34 | ODPB_BACST |
| TKN1_PHYFU | 1 | 11   | 52.5 | 32 | 32 | TKN1_PHYFU |
| IBP1_BOVIN | 1 | 263  | 52.5 | 32 | 32 | IBP1_BOVIN |
| TRPA_THETH | 1 | 271  | 52.5 | 32 | 32 | TRPA_THETH |
| YDHH_HAEIN | 1 | 382  | 52.5 | 32 | 32 | YDHH_HAEIN |
| GCH2_CHLPN | 1 | 418  | 52.5 | 32 | 32 | GCH2_CHLPN |
| RHO_BORBU  | 1 | 419  | 52.5 | 32 | 32 | RHO_BORBU  |
| RHO_PSEFL  | 1 | 419  | 52.5 | 32 | 32 | RHO_PSEFL  |
| RHO_THEMEA | 1 | 427  | 52.5 | 32 | 32 | RHO_THEMEA |
| MM12_MOUSE | 1 | 462  | 52.5 | 32 | 32 | MM12_MOUSE |
| MM12_RAT   | 1 | 465  | 52.5 | 32 | 32 | MM12_RAT   |
| MM01_PIG   | 1 | 469  | 52.5 | 32 | 32 | MM01_PIG   |
| SYK_BACST  | 1 | 494  | 52.5 | 32 | 32 | SYK_BACST  |
| RHO_TREPA  | 1 | 519  | 52.5 | 32 | 32 | RHO_TREPA  |
| RAI2_MOUSE | 1 | 529  | 52.5 | 32 | 32 | RAI2_MOUSE |
| RAI2_HUMAN | 1 | 530  | 52.5 | 32 | 32 | RAI2_HUMAN |
| CAC3_DROME | 1 | 535  | 52.5 | 32 | 32 | CAC3_DROME |
| PGH2_RABIT | 1 | 604  | 52.5 | 32 | 32 | PGH2_RABIT |
| DAB2_MOUSE | 1 | 766  | 52.5 | 32 | 32 | DAB2_MOUSE |
| DSC3_BOVIN | 1 | 896  | 52.5 | 32 | 32 | DSC3_BOVIN |
| STA2_MOUSE | 1 | 923  | 52.5 | 32 | 32 | STA2_MOUSE |
| Y124_METUA | 1 | 1075 | 52.5 | 32 | 32 | Y124_METUA |
| RPOB_HETCA | 1 | 1116 | 52.5 | 32 | 32 | RPOB_HETCA |
| FM14_MOUSE | 1 | 1206 | 52.5 | 32 | 32 | FM14_MOUSE |
| FMN1_MOUSE | 1 | 1213 | 52.5 | 32 | 32 | FMN1_MOUSE |
| VIT6_CAEEL | 1 | 1468 | 52.5 | 32 | 32 | VIT6_CAEEL |
| SC16_YEAST | 1 | 1651 | 52.5 | 32 | 32 | SC16_YEAST |
| TKN2_UPERU | 1 | 2194 | 52.5 | 31 | 31 | TKN2_UPERU |
| MOT2_MERUN | 1 | 71   | 50.8 | 31 | 31 | MOT2_MERUN |
| FTRV_MAIZE | 1 | 97   | 50.8 | 31 | 31 | FTRV_MAIZE |
| YD43_MYCLE | 1 | 126  | 50.8 | 31 | 31 | YD43_MYCLE |
| RS16_YEAST | 1 | 142  | 50.8 | 31 | 31 | RS16_YEAST |
| PR39_PIG   | 1 | 172  | 50.8 | 31 | 31 | PR39_PIG   |
| TBP_ARCFU  | 1 | 183  | 50.8 | 31 | 31 | TBP_ARCFU  |
| MOVV_TOML2 | 1 | 264  | 50.8 | 31 | 31 | MOVV_TOML2 |
| MOVV_TOML2 | 1 | 264  | 50.8 | 31 | 31 | MOVV_TOML2 |
| MOVV_TOML2 | 1 | 264  | 50.8 | 31 | 31 | MOVV_TOML2 |
| YOTB_CAEEL | 1 | 266  | 50.8 | 31 | 31 | YOTB_CAEEL |
| THIM_PASMU | 1 | 267  | 50.8 | 31 | 31 | THIM_PASMU |
| CCHL_YEAST | 1 | 269  | 50.8 | 31 | 31 | CCHL_YEAST |
| RRP1_YEAST | 1 | 278  | 50.8 | 31 | 31 | RRP1_YEAST |
| LST_HAEIN  | 1 | 283  | 50.8 | 31 | 31 | LST_HAEIN  |
| NUGM_NEUCR | 1 | 304  | 50.8 | 31 | 31 | NUGM_NEUCR |
| VPRT_SMRVH | 1 | 322  | 50.8 | 31 | 31 | VPRT_SMRVH |
| PUR5_METHH | 1 | 338  | 50.8 | 31 | 31 | PUR5_METHH |
| PYRD_PASMU | 1 | 339  | 50.8 | 31 | 31 | PYRD_PASMU |
| CPXE_STRGO | 1 | 405  | 50.8 | 31 | 31 | CPXE_STRGO |
| TIG_MYCTO  | 1 | 466  | 50.8 | 31 | 31 | TIG_MYCTO  |
| TRE2_SYNV3 | 1 | 485  | 50.8 | 31 | 31 | TRE2_SYNV3 |
| EXON_HSV6U | 1 | 488  | 50.8 | 31 | 31 | EXON_HSV6U |
| EXON_HSV6Z | 1 | 488  | 50.8 | 31 | 31 | EXON_HSV6Z |

107 31 50.8 488 1 SYK\_MYCHO  
 108 31 50.8 509 1 CPV1\_BRARE  
 109 31 50.8 510 1 DHAF\_VIBHA  
 110 31 50.8 546 1 YTE4\_CAEEL  
 111 31 50.8 558 1 GPC1\_RAT  
 112 31 50.8 660 1 MM02\_HUMAN  
 113 31 50.8 662 1 MM02\_MOUSE  
 114 31 50.8 662 1 MM02\_RABIT  
 115 31 50.8 662 1 MM02\_RAT  
 116 31 50.8 663 1 MM02\_CHICK  
 117 31 50.8 685 1 SNWA\_DICTDI  
 118 31 50.8 687 1 AK48\_RAT  
 119 31 50.8 692 1 AK48\_HUMAN  
 120 31 50.8 704 1 PNP\_BACSU  
 121 31 50.8 768 1 LIPS\_RAT  
 122 31 50.8 864 1 E78A\_DROME  
 123 31 50.8 922 1 YB1C\_SCHPO  
 124 31 50.8 956 1 PMAB\_ARATH  
 125 31 50.8 1629 1 AT59\_HUMAN  
 126 31 50.8 3712 1 ACVS\_CEPAC  
 127 30.5 50.0 2261 1 ABC1\_MOUSE

## ALIGNMENTS

RESULT 1  
 ID TKNA\_HORSE STANDARD; PRT; 11 AA.  
 AC P01250;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE SUBSTANCE P.  
 GN TAC1 OR NKNA OR TAC2 OR NKA.  
 OS Equus caballus (Horse), and Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796, 10141;  
 RN [1]  
 RP SEQUENCE.  
 RC SPECIES=Horse;  
 RA Studer R.O., Trzeciak A., Lergler W.;  
 RT "Isolation and amino-acid sequence of substance P from horse  
 intestine.";  
 RL Helv. Chim. Acta 56:860-866(1973).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=C. porcellus;  
 RX MEDLINE=90044685; PubMed=2478925;  
 RA Murphy R.;  
 RT "Primary amino acid sequence of guinea-pig substance P.";  
 RL Neuropeptides 14:105-110(1989).  
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOLVE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR PIR; A01558; SPHO.  
 DR PIR; A0654; A60654.  
 DR InterPro: IPR003580; Protachykinin.  
 DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR SMART; SM00203; TK; 1.  
 DR PROSITE; PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation; Neurotransmitter.  
 FT MOD\_RES 11 11 AMIDATION  
 FT SEQUENCE 11 AA; 1349 MW; 3E757FE3C9D6C6C7 CRC64;

Query Match 100.0%; Score 61; DB 1; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOFFFLM 11  
 Db 1 RPKPQOFFFLM 11  
 RESULT 2  
 ID TKNL\_RABIT STANDARD; PRT; 115 AA.  
 AC P41540;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROTHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
 (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE GAMMA; C-TERMINAL  
 FLANKING PEPTIDE].  
 GN TAC1 OR NKNA OR TAC2 OR NKA.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=93371392; PubMed=8363593;  
 RA Maegert H.J., Heitland A., Rose M., Forssmann W.G.;  
 RT "Nucleotide sequence of the rabbit gamma-preprotachykinin I CDNA.";  
 RL Biochem. Biophys. Res. Commun. 195:128-131(1993).  
 RN [2]  
 RP SEQUENCE OF 72-92.  
 RA Kage R., McGregor G.P., Thim L., Conlon J.M.;  
 RT "Gamma-neuropeptide K: a peptide isolated from rabbit gut that is  
 derived from gamma-preprotachykinin.";  
 RL Regul. Pept. 18:346-346(1987).  
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOLVE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
 CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X62994; CAA44728.1; -  
 DR PIR; S18922; S18922.  
 DR InterPro: IPR003580; Protachykinin.  
 DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR Prodom; PD005598; Protachykinin; 1.  
 DR SMART; SM00203; TK; 2.  
 DR PROSITE; PS00267; TACHYKININ; 2.  
 KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
 KW Amidation; Alternative splicing; Signal; Neurotransmitter.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT PEPTIDE 20 56 POTENTIAL.  
 FT PEPTIDE 58 68 SUBSTANCE P.  
 FT PEPTIDE 72 92 NEUROPEPTIDE GAMMA.  
 FT PEPTIDE 83 92 NEUROKININ A.  
 FT PEPTIDE 96 111 C-TERMINAL FLANKING PEPTIDE.  
 FT MOD\_RES 68 88 AMIDATION (G-69 PROVIDE AMIDE GROUP).  
 FT MOD\_RES 92 92 AMIDATION (G-93 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 115 AA; 13370 MW; 5EC76F7C9B10E1C6 CRC64;

Query Match 100.0%; Score 61; DB 1; Length 115;  
 Best Local Similarity 100.0%; Pred. No. 0.00014;



Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11  
|||||

Db 58 RPKPQQFFGLM 68

## RESULT 3

TKNL\_HUMAN STANDARD; PRT; 129 AA.  
AC P20366; Q00072; O60600; O60601;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROTACHYKININ 1 PRECURSOR (PT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE  
DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
GN TACI OR NKNA OR TAC2 OR NKA.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM BETA).  
RX MEDLINE=87030957; PubMed=3770210;  
RA Harmar A.J., Armstrong A., Pascall J.C., Chapman K., Rosie R.,  
RA Curtis A., Going J., Edwards C.R.W., Fink G.;  
RT "cDNA sequence of human beta-preprotachykinin, the common precursor  
RT to substance P and neurokinin A.";  
RL FEBS Lett. 208:67-72(1986).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM BETA).  
RC TISSUE=Brain;  
RA Tan A., Foo H.P.;  
RN Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 36-122 FROM N.A. (ISOFORMS BETA AND GAMMA).  
RX MEDLINE=91209287; PubMed=1708336;  
RA Chikakata C., Brackmann B., Hunt N., Davidoff M., Schulze W.,  
RA Ivell R.;  
RT "Tachykinin (substance-P) gene expression in Leydig cells of the  
RT human and mouse testis.";  
RL Endocrinology 128:2441-2448(1991).  
RN [4]  
RP SEQUENCE OF 98-107.  
RX MEDLINE=87275962; PubMed=3038549;  
RA Theodorsson-Norheim E., Joernvall H., Andersson M., Norheim I.,  
RA Oberg K., Jacobsson G.;  
RT "Isolation and characterization of neurokinin A, neurokinin A(3-10)  
RT and neurokinin A(4-10) from a neutral water extract of a metastatic  
RT ileal carcinoma tumour.";  
RL Eur. J. Biochem. 166:693-697(1987).  
RN [5]  
RP SEQUENCE OF 36-118 FROM N.A. (ISOFORM ALPHA).  
RX TISSUE=Blood, and Brain;  
RA Lai J.P., Douglas S.D., Rappaport E., Wu J.M., Ho W.Z.;  
RT "Identification of a delta isoform of preprotachykinin mRNA in human  
RT mononuclear phagocytes and lymphocytes.";  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE OF 111-126.  
RX TISSUE=Adrenal medulla;  
RA MEDLINE=91133994; PubMed=2284201;  
RA McGregor G.P., Conlon J.M.;  
RT "Characterization of the C-terminal flanking peptide of human  
RT beta-preprotachykinin.";  
RL Peptides 11:907-910(1990).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SEROTONINERGIC, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),

GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
-1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X54469; CAA38351.1; -  
DR EMBL; U37529; AAA79195.1; -  
DR EMBL; M68906; AAA60159.1; -  
DR EMBL; M68907; AAA60160.1; -  
DR EMBL; AF050656; AAC15702.1; -  
DR EMBL; AF050658; AAC15704.1; -  
DR PIR; A24805; A24805.  
DR PIR; S00069; S00069.  
DR MIN; 162320; -  
DR InterPro; IPR003580; Protachykinin.  
DR InterPro; IPR002040; Tachykinin.  
DR Pfam; PF02202; tachykinin; 1.  
DR ProDom; PD005598; Protachykinin; 1.  
DR SMART; SM00203; TK; 2.  
DR PROSITE; PS00267; TACHYKININ; 2.  
KW Tachykinin; Neuroptide; Cleavage on pair of basic residues;  
KW Amidation; Alternative splicing; Signal; Neurotransmitter.  
FT SIGNAL 1 19  
FT PROPEP 20 56  
FT PEPTIDE 58 68  
FT PEPTIDE 72 107  
FT PEPTIDE 72 73  
FT PEPTIDE 89 107  
FT PEPTIDE 98 107  
FT PEPTIDE 111 126  
FT MOD\_RES 68 68  
FT MOD\_RES 107 107  
FT VARSPIC 74 88  
FT VARSPIC 97 114  
FT VARSPIC 115 115  
FT VARSPIC 87 87  
FT CONFLICT 87 87  
SQ SEQUENCE 129 AA; 15003 MW; 51412C1692368DE4 CRC64;  
L -> P (IN REF. 4).  
MISSING (IN ISOFORM ALPHA AND ISOFORM  
DELTA).  
MISSING (IN ISOFORM ALPHA AND ISOFORM  
DELTA).  
V -> M (IN ISOFORM ALPHA AND ISOFORM  
DELTA).  
L -> P (IN REF. 4).  
Query Match 100.0%; Score 61; DB 1; Length 129;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 RPKPQQFFGLM 11  
|||||  
Db 58 RPKPQQFFGLM 68  
RESULT 4  
TKNL\_BOVIN STANDARD; PRT; 130 AA.  
AC P01289; P01291; P04091; P20773;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROTACHYKININ 1 PRECURSOR (PT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE  
DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
GN TACI OR NKNA OR TAC2 OR NKA.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;

[1]  
SEQUENCE FROM N.A. (ISOFORM BETA).  
MEDLINE=85086245; PubMed=6083453;  
Nawa H., Kotani H., Nakanishi S.  
"Tissue-specific generation of two preprotachykinin mRNAs from one  
gene by alternative RNA splicing.";  
Nature 312:729-734(1984).  
[2]  
SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).  
MEDLINE=84039802; PubMed=6195531;  
Nawa H., Hirose T., Takashima H., Inayama S., Nakanishi S.;  
"Nucleotide sequences of cloned cDNAs for two types of bovine brain  
substance P precursor.";  
Nature 306:32-36(1983).  
[3]  
SEQUENCE OF 36-122 FROM N.A. (ISOFORMS BETA AND GAMMA).  
TISSUE-HYPOTHALAMUS;  
MEDLINE=91209287; PubMed=1708336;  
Chiwakata C., Brackmann B., Hunt N., Davidoff M., Schulze W.,  
Ivell R.;  
"Tachykinin (substance-P) gene expression in Leydig cells of the  
human and mouse testis.";  
Endocrinology 128:2441-2448(1991).  
-1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
MUSCLES.  
-1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
-1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X00075; CAA24939.1; -  
DR EMBL; X00075; CAA24940.1; -  
DR EMBL; X00075; CAA24941.1; -  
DR EMBL; X00076; CAA24942.1; -  
DR EMBL; X00076; CAA24943.1; ALT\_SEQ.  
DR EMBL; X02351; CAA26206.1; -  
DR EMBL; X01396; CAA26206.1; JOINED.  
DR EMBL; X01397; CAA26206.1; JOINED.  
DR EMBL; X01398; CAA26206.1; JOINED.  
DR EMBL; X01399; CAA26206.1; JOINED.  
DR EMBL; X01400; CAA26206.1; JOINED.  
DR EMBL; M68911; AAA30724.1; -  
DR EMBL; M68912; AAA30725.1; -  
DR PIR; A01557; SPBOA.  
DR PIR; A01559; SPBOB.  
DR PIR; A05093; A05093.  
DR PIR; B25067; B25067.  
DR InterPro; IPR003580; Protachykinin.  
DR InterPro; IPR002040; Tachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR ProDom; PD005598; Protachykinin; 1.  
DR SMART; SM00203; TK; 2.  
DR PROSITE; PS00267; TACHYKININ; 2.  
Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
Amidation; Alternative splicing; Signal; Neurotransmitter.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 56 POTENTIAL.  
FT PEPTIDE 58 68 SUBSTANCE P.  
FT PEPTIDE 72 107 NEUROPEPTIDE K.  
FT PEPTIDE 72 73 NEUROPEPTIDE GAMMA 1ST PART.  
FT PEPTIDE 89 107 NEUROPEPTIDE GAMMA 2ND PART.  
FT PEPTIDE 98 107 NEUROKININ A.  
FT PEPTIDE 111 126 C-TERMINAL FLANKING PEPTIDE (POTENTIAL).  
FT MOD\_RES 68 68 AMIDATION (G-69 PROVIDE AMIDE GROUP).  
-----

FT MOD\_RES 107 107 AMIDATION (G-108 PROVIDE AMIDE GROUP).  
FT VARSPLIC 74 88 MISSING (IN ISOFORM GAMMA AND ISOFORM  
DELTA).  
FT VARSPLIC 97 114 MISSING (IN ISOFORM ALPHA AND ISOFORM  
DELTA).  
FT VARSPLIC 115 115 V -> M (IN ISOFORM ALPHA AND ISOFORM  
DELTA).  
FT CONFLICT 121 121 V -> A (IN REF. 3).  
SQ SEQUENCE 130 AA; 15076 MW; CE2A28572305DEB7 CRC64;  
  
Query Match 100.0%; Score 61; DB 1; Length 130;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RPKPQQFFGLM 11  
Db 58 RPKPQQFFGLM 68  
|||||  
  
RESULT 5  
TKNL\_MESAU  
ID TKNL\_MESAU STANDARD; PRT; 130 AA.  
AC Q60541; P49110;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROTACHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
(NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE  
DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
GN TAC1 OR NKNA OR TAC2 OR NKA.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus.  
OX NCBI\_TaxID=10036;  
RN [1]  
SEQUENCE FROM N.A. (ISOFORMS BETA AND GAMMA).  
R STRAIN-AURA; TISSUE-Brain;  
RA Heiland A., Kruhoffer M., Juerger Maegert H.J., Forssmann W.G.;  
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
MUSCLES.  
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; X80662; CAA56691.1; -  
DR EMBL; X80663; CAA56692.1; -  
DR InterPro; IPR003580; Protachykinin.  
DR InterPro; IPR002040; Tachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR ProDom; PD005598; Protachykinin; 1.  
DR SMART; SM00203; TK; 2.  
DR PROSITE; PS00267; TACHYKININ; 2.  
Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
Amidation; Alternative splicing; Signal; Neurotransmitter.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 56 POTENTIAL.  
FT PEPTIDE 58 68 SUBSTANCE P.  
FT PEPTIDE 72 107 NEUROPEPTIDE K.  
FT PEPTIDE 72 73 NEUROPEPTIDE GAMMA 1ST PART.  
FT PEPTIDE 89 107 NEUROPEPTIDE GAMMA 2ND PART.

FT PEPTIDE 98 107 NEUROKININ A.  
FT PEPTIDE 111 126 C-TERMINAL FLANKING PEPTIDE (POTENTIAL).  
FT MOD\_RES 68 AMIDATION (G-69 PROVIDE AMIDE GROUP).  
FT MOD\_RES 107 107 AMIDATION (G-108 PROVIDE AMIDE GROUP).  
FT VARSPLIC 74 88 MISSING (IN ISOFORM GAMMA).  
SQ SEQUENCE 130 AA; 14907 MW; CC92E9371A646F2E CRC64;

Query Match 100.0%; Score 61; DB 1; Length 130;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPKPQOFFGLM 11  
|||||  
DB 58 RPKPQOFFGLM 68

RESULT 6  
TKNL\_MOUSE STANDARD; PRT; 130 AA.  
AC P41539; Q00073;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROTACHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE  
DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
GN TAC1 OR NKNA OR TAC2 OR NKA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM BETA).  
RC STRAIN=ICR; TISSUE=Brain;  
RA Kako K., Muneoka E., Hosaka M., Murakami K., Nakayama K.;  
RT "Cloning and sequence analysis of mouse cDNAs encoding  
preprotachykinin A and B."  
RL Biomed. Res. 14:253-259(1993).  
RN [2]  
RP SEQUENCE OF 36-122 FROM N.A. (ISOFORMS BETA AND GAMMA).  
RC TISSUE=Brain;  
RX MEDLINE=91209287; PubMed=1708336;  
RA Chiwakata C., Brackmann B., Hunt N., Davidoff M., Schulze W.,  
RA Ivell R.;  
RT "Tachykinin (substance-P) gene expression in Leydig cells of the  
human and mouse testis."  
RL Endocrinology 128:2441-2448(1991).  
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
MUSCLES.  
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D17584; BAA04508.1; -  
DR EMBL; M68908; AAA39969.1; -  
DR EMBL; M68909; AAA39970.1; -  
DR MGD; MGI:98474; Tac1.  
DR InterPro; IPR003580; Protachykinin.  
DR InterPro; IPR002040; Tachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR ProDom; PD005598; Protachykinin; 1.  
DR SMART; SM00203; TK; 2.

DR PROSITE; PS00267; TACHYKININ; 2.  
KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
KW Amidation; Alternative splicing; Signal; Neurotransmitter.  
FT SIGNAL 1 19 POTENTIAL.  
FT PROPEP 20 56 POTENTIAL.  
FT PEPTIDE 58 68 SUBSTANCE P.  
FT PEPTIDE 72 107 NEUROPEPTIDE K.  
FT PEPTIDE 72 73 NEUROPEPTIDE GAMMA 1ST PART.  
FT PEPTIDE 89 107 NEUROPEPTIDE GAMMA 2ND PART.  
FT PEPTIDE 98 107 NEUROKININ A.  
FT PEPTIDE 111 126 C-TERMINAL FLANKING PEPTIDE (POTENTIAL).  
FT MOD\_RES 68 68 AMIDATION (G-69 PROVIDE AMIDE GROUP).  
FT MOD\_RES 107 107 AMIDATION (G-108 PROVIDE AMIDE GROUP).  
FT VARSPLIC 74 88 MISSING (IN ISOFORM GAMMA).  
SQ SEQUENCE 130 AA; 15045 MW; 7BE8DA15FDE72FF8 CRC64;

Query Match 100.0%; Score 61; DB 1; Length 130;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPKPQOFFGLM 11  
|||||  
DB 58 RPKPQOFFGLM 68

RESULT 7  
TKNL\_RAT STANDARD; PRT; 130 AA.  
AC P06767; P08856; P08857; P22356;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROTACHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE  
DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
GN TAC1 OR NKNA OR TAC2 OR NKA.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA AND GAMMA).  
RX MEDLINE=90331040; PubMed=1695945;  
RA Carter M.S., Krause J.E.;  
RT "Structure, expression, and some regulatory mechanisms of the rat  
preprotachykinin gene encoding substance P, neurokinin A,  
neuropeptide K, and neuropeptide gamma."  
RL J. Neurosci. 10:2203-2214(1990).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA AND GAMMA).  
RX MEDLINE=87118268; PubMed=2433692;  
RA Krause J.E., Chirgwin J.M., Carter M.S., Xu Z.S., Hershey A.D.;  
RT "Three rat preprotachykinin mRNAs encode the neuropeptides substance  
P and neurokinin A."  
RL Proc. Natl. Acad. Sci. U.S.A. 84:881-885(1987).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM GAMMA).  
RX MEDLINE=87025808; PubMed=2429656;  
RA Kawaguchi Y., Hoshimaru M., Nawa H., Nakanishi S.;  
RT "Sequence analysis of cloned cDNA for rat substance P precursor:  
existence of a third substance P precursor."  
RL Biochem. Biophys. Res. Commun. 139:1040-1046(1986).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM DELTA).  
RC TISSUE=Dorsal root ganglion;  
RX MEDLINE=91085565; PubMed=1702066;  
RA Hamar A.J., Hyde V., Chapman K.E.;  
RT "Identification and cDNA sequence of delta-preprotachykinin, a fourth  
splicing variant of the rat substance P precursor."  
RL FEBS Lett. 275:22-24(1990).  
RN [5]  
RP SEQUENCE OF 1-41 FROM N.A.

RX MEDLINE=93192337; PubMed=8448217;  
RA Chapman K.E., Lyons V., Harmar A.J.;  
RT "The sequence of 5' flanking DNA from the rat preprotachykinin gene;  
RL analysis of putative transcription factor binding sites.";  
RL Biochem. Biophys. Acta 1172:361-363(1993).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M34162; AAA41926.1; -  
CC EMBL; M34159; AAA41926.1; JOINED.  
CC EMBL; M34160; AAA41926.1; JOINED.  
CC EMBL; M34161; AAA41926.1; JOINED.  
CC EMBL; M34184; AAA41925.1; -  
CC EMBL; M34183; AAA41929.1; -  
CC EMBL; M15191; AAA41928.1; -  
CC EMBL; M14312; AAA41927.1; -  
CC EMBL; L07328; AAA41924.1; -  
CC EMBL; X56306; CAA39752.1; -  
CC PIR; A26590; A26590.  
CC PIR; B26590; B26590.  
CC PIR; C26590; C26590.  
CC PIR; A37163; A37163.  
CC PIR; S12958; S12958.  
CC InterPro; IPR003580; Protachykinin.  
CC InterPro; IPR002040; Tachykinin.  
CC Pfam; PF02202; Tachykinin; 1.  
CC ProDom; PD005598; Protachykinin; 1.  
CC SMART; SM00203; TK; 2.  
CC PROSITE; PS00267; TACHYKININ; 2.  
CC Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
CC Amidation; Alternative splicing; Signal; Neurotransmitter.  
FT SIGNAL 1 19  
FT PROPEP 20 56  
FT PEPTIDE 58 68  
FT PEPTIDE 72 107  
FT PEPTIDE 72 73  
FT PEPTIDE 89 107  
FT PEPTIDE 98 107  
FT PEPTIDE 111 126  
FT PEPTIDE 111 126  
FT MOD\_RES 68 68  
FT MOD\_RES 107 107  
FT VARSPPLIC 74 88  
FT VARSPPLIC 97 114  
FT VARSPPLIC 115 115  
FT VARSPPLIC 115 115  
FT V->M (IN ISOFORM ALPHA AND ISOFORM  
FT DELTA).  
FT DELTA).  
SQ SEQUENCE 130 AA; 15001 MW; B22EFE860DCD75A CRC64;

Query Match 100.0%; Score 61; DB 1; Length 130;  
Best Local Similarity 100.0%; Pred. No. 0.00016;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11  
|||||

Db 58 RPKPQQFFGLM 68  
|||||

RESULT 8

TKNA\_CHICK STANDARD; PRT; 11 AA.  
ID TKNA\_CHICK  
AC P19850;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE SUBSTANCE P.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Intestine;  
RX MEDLINE=88204263; PubMed=2452461;  
RA Conlon J.M., Katsoulis S., Schmidt W.E., Thim L.;  
RT "[Arg3]substance P and neurokinin A from chicken small intestine.";  
RL Regul. Pept. 20:171-180(1988).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
CC PIR; JN0023; JN0023.  
CC InterPro; IPR003580; Protachykinin.  
CC InterPro; IPR002040; Tachykinin.  
CC Pfam; PF02202; Tachykinin; 1.  
CC SMART; SM00203; TK; 1.  
CC PROSITE; PS00267; TACHYKININ; 1.  
CC Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 11 11  
FT MOD\_RES 11 11  
FT SEQUENCE 11 AA; 1377 MW; 21487FE3C9D6C6C7 CRC64;  
Query Match 95.1%; Score 58; DB 1; Length 11;  
Best Local Similarity 90.9%; Pred. No. 5e-05;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 RPKPQQFFGLM 11  
|||||  
Db 1 RPRPQQFFGLM 11  
|||||  
RESULT 9  
TKNA\_ONCMY STANDARD; PRT; 11 AA.  
ID TKNA\_ONCMY  
AC P28499;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE SUBSTANCE P.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Brain;  
RX MEDLINE=92298992; PubMed=1376687;  
RA Jensen J., Conlon J.M.;  
RT "Substance-P-related and neurokinin-A-related peptides from the brain  
RT of the cod and trout.";  
RL Eur. J. Biochem. 206:659-664(1992).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
CC PIR; S23307; S23307.  
CC PIR; S23308; S23308.  
CC InterPro; IPR003580; Protachykinin.

DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation; Neurotransmitter.  
 FT MOD\_RES 11 11 AMIDATION (BY SIMILARITY).  
 SQ SEQUENCE 11 AA; 1358 MW; 214860DEC9D6D1F7 CRC64;

Query Match 82.0%; Score 50; DB 1; Length 11;  
 Best Local Similarity 72.7%; Pred. No. 0.0015;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11  
 : : | | | | |  
 Db 1 KPRPHQFFGLM 11

RESULT 10  
 TKNA\_GADMO STANDARD; PRT; 11 AA.  
 AC P28498;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE SUBSTANCE P.  
 OS Gadus morhua (Atlantic cod).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadoidei; Gadidae;  
 OC Gadus.  
 OX NCBI\_TaxID=8049;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=92298992; PubMed=1376687;  
 RA Jensen J., Conlon J.M.;  
 RT "Substance P-related and neurokinin-A-related peptides from the brain  
 of the cod and trout."  
 RL Eur. J. Biochem. 206:659-664(1992).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR InterPro: IPR003580; Protachykinin.  
 DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation; Neurotransmitter.  
 FT MOD\_RES 11 11 AMIDATION (BY SIMILARITY).  
 SQ SEQUENCE 11 AA; 1315 MW; 214860D759D6C6C7 CRC64;

Query Match 80.3%; Score 49; DB 1; Length 11;  
 Best Local Similarity 72.7%; Pred. No. 0.0023;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11  
 : : | | | | |  
 Db 1 KPRPQQFFGLM 11

RESULT 11  
 TKNA\_SCYCA STANDARD; PRT; 11 AA.  
 AC P41333;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE SUBSTANCE P.  
 OS Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;  
 OC Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;  
 OC Scyliorhinidae; Scyliorhinus.  
 OX NCBI\_TaxID=7830;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Brain;  
 RX MEDLINE=93292508; PubMed=7685693;  
 RA Waugh D., Wang Y., Hazon N., Balmert R.J., Conlon J.M.;  
 RT "Primary structures and biological activities of substance P-related  
 peptides from the brain of the dogfish, Scyliorhinus canicula.";  
 RL Eur. J. Biochem. 214:469-474(1993).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR PIR: S33300; S33300.  
 DR InterPro: IPR003580; Protachykinin.  
 DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation; Neurotransmitter.  
 FT MOD\_RES 11 11 AMIDATION.  
 SQ SEQUENCE 11 AA; 1278 MW; 214860DEC9D6D867 CRC64;

Query Match 78.7%; Score 48; DB 1; Length 11;  
 Best Local Similarity 72.7%; Pred. No. 0.0036;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11  
 : : | | | | |  
 Db 1 KPRPQQFFGLM 11

RESULT 12  
 TKNA\_PSEGU STANDARD; PRT; 11 AA.  
 AC P42989;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE SUBSTANCE P-LIKE PEPTIDE I (PG-SPI).  
 OS Pseudophryne guentheri (Frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
 OC Pseudophryne.  
 OX NCBI\_TaxID=30349;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=Skin;  
 RX MEDLINE=90287814; PubMed=2356157;  
 RA Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,  
 RA Roberts J.D., Melchiorri P., Erspamer V.;  
 RT "Six novel tachykinin- and bombesin-related peptides from the skin of  
 the Australian frog Pseudophryne guentheri.";  
 RL Peptides 11:299-304(1990).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR PIR: E60409; E60409.  
 DR InterPro: IPR003580; Protachykinin.  
 DR InterPro: IPR002040; Tachykinin.  
 DR Pfam: PF02202; Tachykinin; 1.  
 DR SMART: SM00203; TK; 1.  
 DR PROSITE: PS00267; TACHYKININ; 1.  
 KW Tachykinin; Neuropeptide; Amidation.  
 FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
 FT MOD\_RES 11 11 AMIDATION.

```

RN RP SEQUENCE.
RA Yasuhara T., Nakajima T., Erspamer G.F., Erspamer V.;
RT "New tachykinins, Glu2, Pro5-kassinin (hylambates-kassinin) and
RT hylambatin, in the skin of the African rhacophorid frog Hylambates
RT maculatus.";
CC Biomed. Res. 2:613-617(1981).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC PIR: S10059; S10059.
CC DR InterPro: IPR003580; Protachykinin.
CC DR InterPro: IPR002040; Tachykinin.
CC DR Pfam: PF02202; Tachykinin; 1.
CC DR SMART: SM00203; TK; 1.
CC DR PROSITE: PS00267; TACHYKININ; 1.
CC DR Tachykinin; Neuropeptide; Amidation; Amphibian skin.
CC KW MOD_RES 12 12 AMIDATION.
CC SQ SEQUENCE 12 AA; 1376 MW; 3E756D279DD6DAB7 CRC64;

Query Match 72.1%; Score 44; DB 1; Length 12;
Best Local Similarity 80.0%; Pred. No. 0.022;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PKPQQFFGLM 11
Db 3 PKPDQFVGLM 12

RESULT 15
TKNA_RANRI
ID TKNA_RANRI STANDARD; PRT; 11 AA.
AC P29207;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RANAKININ (SUBSTANCE-P-RELATED PEPTIDE).
OS Rana ridibunda (Laughing frog) (Marsh frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Rana.
OX NCBI_TaxID=8406;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RX MEDLINE=32044543; PubMed=1658233;
RA O'Harte F., Burcher E., Lovas S., Smith D.D., Vaudry H., Conlon J.M.;
RT "Ranakinin: a novel NK1 tachykinin receptor agonist isolated with
RT neurokinin B from the brain of the frog Rana ridibunda.";
RL J. Neurochem. 57:2086-2091(1991).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC DR InterPro: IPR003580; Protachykinin.
CC DR InterPro: IPR002040; Tachykinin.
CC DR Pfam: PF02202; Tachykinin; 1.
CC DR SMART: SM00203; TK; 1.
CC DR PROSITE: PS00267; TACHYKININ; 1.
CC DR Tachykinin; Neuropeptide; Amidation.
CC KW MOD_RES 11 11 AMIDATION.
CC SQ SEQUENCE 11 AA; 1352 MW; 3A2460CC59D40B07 CRC64;

Query Match 70.5%; Score 43; DB 1; Length 11;
Best Local Similarity 54.8%; Pred. No. 0.03;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11

```

Db 1 KPNPERFYGLM 11

RESULT 16

TKNA\_RANCA STANDARD; PRT; 11 AA.

AC P22688;

DT 01-AUG-1991 (Rel. 19, Created)

DT 01-AUG-1991 (Rel. 19, Last sequence update)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DE RANATACHYKININ A (RHK A).

OS Rana catesbeiana (Bull frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoides; Ranidae; Rana.

OX NCBI\_TaxID=8400;

RN [1]

RP SEQUENCE, AND SYNTHESIS.

RC TISSUE=Brain, and Intestine;

RX MEDLINE=91254337; PubMed=2043143;

RA Kozawa H., Hino J., Minamino N., Kangawa K., Matsuo H.;

RT "Isolation of four novel tachykinins from frog (Rana catesbeiana) brain and intestine.";

RL Biochem. Biophys. Res. Commun. 177:588-595(1991).

RN [2]

RP SEQUENCE.

RC TISSUE=Intestine;

RX MEDLINE=94023216; PubMed=8210506;

RA Kangawa K., Kozawa H., Hino J., Minamino N., Matsuo H.;

RT "Four novel tachykinins in frog (Rana catesbeiana) brain and intestine.";

RL Regul. Pept. 46:81-88(1993).

CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS, EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH MUSCLES.

CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.

DR PIR; JEO426; JEO426.

DR PIR; A61033; A61033.

DR InterPro; IPR003580; Protachykinin.

DR InterPro; IPR002040; Tachykinin.

DR Pfam; PF02202; Tachykinin; 1.

DR SMART; SM00203; TK; 1.

DR PROSITE; PS00267; TACHYKININ; 1.

KW Tachykinin; Neuropeptide; Amidation.

FT MOD\_RES 11 11 AMIDATION.

SQ SEQUENCE 11 AA; 1311 MW; 200D60CC59D40AB7 CRC64;

Query Match 67.2%; Score 41; DB 1; Length 11;  
Best Local Similarity 54.5%; Pred. No. 0.071;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQOFFGLM 11  
:|:|:|:|

Db 1 KPSDPFRFYGLM 11

RESULT 17

TRX\_DROVI STANDARD; PRT; 3828 AA.

AC Q24742;

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE TRITHORAX PROTEIN.

GN TRX.

OS Drosophila virilis (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI\_TaxID=7244;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96100387; PubMed=85551104;

RA Tillib S., Sedkov Y., Mizrokhi L., Mazo A.;

RT "Conservation of structure and expression of the trithorax gene between Drosophila virilis and Drosophila melanogaster.";

RL Mech. Dev. 53:113-122(1995).

CC -!- FUNCTION: FUNCTIONS IN SEGMENT DETERMINATION THROUGH INTERACTION WITH GENES OF BITHORAX (BX-C) AND ANTENNAPEDIA (ANT-X) COMPLEXES.

CC IT CAN BEHAVE AS AN ACTIVATOR OF BX-C.

CC -!- SUBCELLULAR LOCATION: NUCLEAR.

CC -!- SIMILARITY: BELONGS TO THE TRITHORAX FAMILY OF TRANSCRIPTION FACTORS.

CC -!- SIMILARITY: CONTAINS 1 'SET' DOMAIN.

CC -!- SIMILARITY: CONTAINS 5 PHD-TYPE ZINC FINGERS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; Z50038; CAA90349.1; -

DR HSSP; P04002; IWFA.

DR Flybase; FBgn0014844; Dvir\trix.

DR InterPro; IPR003889; FyricH\_C.

DR InterPro; IPR003888; FyricH\_N.

DR InterPro; IPR001965; PHD.

DR InterPro; IPR003616; PostSET.

DR InterPro; IPR001214; SET.

DR InterPro; IPR001841; Znf\_ring.

DR InterPro; IPR001628; znf\_C4.

DR Pfam; PF00628; PHD; 2.

DR Pfam; PF00856; SET; 1.

DR SMART; SM00542; FYRC; 1.

DR SMART; SM00541; FYRN; 1.

DR SMART; SM00249; PHD; 4.

DR SMART; SM00508; PostSET; 1.

DR SMART; SM00184; Ring; 2.

DR SMART; SM00317; SET; 1.

DR SMART; SM00399; Znf\_C4; 1.

DR PROSITE; PS0280; SET; 1.

KW Transcription regulation; Zinc-finger; Metal-binding; DNA-binding; Nuclear protein; Developmental protein; Activator.

FT ZN\_FING 1251 1334 PHD-TYPE 1.

FT ZN\_FING 1335 1380 PHD-TYPE 2.

FT ZN\_FING 1408 1469 PHD-TYPE 3.

FT ZN\_FING 1708 1767 PHD-TYPE 4 (ATYPICAL).

FT ZN\_FING 1768 1818 PHD-TYPE 5 (ATYPICAL).

FT DOMAIN 3701 3810 SET.

FT DOMAIN 28 41 POLY-ALA.

FT DOMAIN 66 71 POLY-ASP.

FT DOMAIN 160 164 POLY-ASP.

FT DOMAIN 173 182 POLY-ALA.

FT DOMAIN 221 228 POLY-GLN.

FT DOMAIN 243 251 POLY-ALA.

FT DOMAIN 253 258 POLY-THR.

FT DOMAIN 292 296 POLY-ALA.

FT DOMAIN 538 546 POLY-ASP.

FT DOMAIN 1072 1075 POLY-GLU.

FT DOMAIN 2483 3271 GLN-RICH.

FT DOMAIN 3333 3339 POLY-ASP.

SQ SEQUENCE 3828 AA; 413721 MW; 32059CF30A3C504 CRC64;

Query Match 67.2%; Score 41; DB 1; Length 3828;  
Best Local Similarity 60.0%; Pred. No. 25;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOFFGL 10  
:|:|:|:|

Db 618 KPKPKNYFGL 627

RESULT 18  
TKN1\_UPERU  
ID TKN1\_UPERU STANDARD; PRT; 11 AA.  
AC P08612;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE UPEROLEIN.  
OS Uperoleia rugosa (Australian leptodactylid frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
OC Uperoleia.  
ON NCBI\_TaxID=8368;  
RX [1]  
RP SEQUENCE.  
RA MEDLINE=75131227; PubMed=1120493;  
RX Anastasi A., Erspamer V., Edean R.;  
RT "Structure of uperolein, a physalaemin-like endecapeptide occurring  
in the skin of Uperoleia rugosa and Uperoleia marmorata.";  
RL Experientia 31:394-395(1975).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR; S07203; S07203.  
DR InterPro; IPR003580; Protachykinin.  
DR InterPro; IPR002040; Tachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR SMART; SM00203; TK; 1.  
DR PROSITE; PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 1 11 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 11 11 AMIDATION.  
SQ SEQUENCE 11 AA; 1252 MW; 32867C3E59CDD457 CRC64;

Query Match 60.7%; Score 37; DB 1; Length 11;  
Best Local Similarity 54.5%; Pred. No. 0.39;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQPFGLM 11  
:| | | | |  
DB 1 QPDPAFYGLM 11

RESULT 19  
TKN2\_PSEGU  
ID TKN2\_PSEGU STANDARD; PRT; 11 AA.  
AC P42987;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE KASSININ-LIKE PEPTIDE K-II (PG-KII).  
OS Pseudophryne guentheri (Frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
OC Pseudophryne.  
ON NCBI\_TaxID=30349;  
RX [1]  
RP SEQUENCE.  
RA TISSUE=Skin;  
RX MEDLINE=90287814; PubMed=2356157;  
RA Sirmaco M., Severini C., de Biase D., Barra D., Bossa F.,  
RA Roberts J.D., Melchiorri P., Erspamer V.;  
RT "Six novel tachykinin- and bombesin-related peptides from the skin of  
the Australian frog Pseudophryne guentheri.";  
RL Peptides 11:299-304(1990).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.

CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR; C60409; C60409.  
DR InterPro; IPR003580; Protachykinin.  
DR InterPro; IPR002040; Tachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR SMART; SM00203; TK; 1.  
DR PROSITE; PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 1 11 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 11 11 AMIDATION.  
SQ SEQUENCE 11 AA; 1246 MW; 3A247C37C9CBIAB7 CRC64;

Query Match 60.7%; Score 37; DB 1; Length 11;  
Best Local Similarity 54.5%; Pred. No. 0.39;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQPFGLM 11  
:| | | | |  
DB 1 QPNPDEFVGLM 11

RESULT 20  
TKN1\_PSEGU  
ID TKN1\_PSEGU STANDARD; PRT; 11 AA.  
AC P42986;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE KASSININ-LIKE PEPTIDE K-I (PG-KI).  
OS Pseudophryne guentheri (Frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;  
OC Pseudophryne.  
ON NCBI\_TaxID=30349;  
RX [1]  
RP SEQUENCE.  
RA TISSUE=Skin;  
RX MEDLINE=90287814; PubMed=2356157;  
RA Sirmaco M., Severini C., de Biase D., Barra D., Bossa F.,  
RA Roberts J.D., Melchiorri P., Erspamer V.;  
RT "Six novel tachykinin- and bombesin-related peptides from the skin of  
the Australian frog Pseudophryne guentheri.";  
RL Peptides 11:299-304(1990).  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR; B60409; B60409.  
DR InterPro; IPR003580; Protachykinin.  
DR InterPro; IPR002040; Tachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR SMART; SM00203; TK; 1.  
DR PROSITE; PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 1 11 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 11 11 AMIDATION.  
SQ SEQUENCE 11 AA; 1269 MW; 3DBA7C37C9CBIAB7 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 11;  
Best Local Similarity 54.5%; Pred. No. 0.6;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQPFGLM 11  
:| | | | |  
DB 1 QPHDPDEFVGLM 11

RESULT 21  
TKN3\_PSEGU  
ID TKN3\_PSEGU STANDARD; PRT; 11 AA.



P42988;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE KASSININ-LIKE PEPTIDE K-III (PG-KIII).  
OS Pseudophryne guentheri (Frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Myobatrachidae;  
OC Neobatrachidae.  
OX NCBI\_TaxID=30349;  
RN [1]  
RP TISSUE=Skin;  
RC MEDLINE=90287814; PubMed=2356157;  
RA Simmaco M., Severini C., de Biase D., Barra D., Bossa F.,  
RA Roberts J.D., Melchiorri P., Erspamer V.;  
RT "Six novel tachykinin- and bombesin-related peptides from the skin of  
the Australian frog Pseudophryne guentheri.";  
RL Peptides 11:299-304(1990).  
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
MUSCLES.  
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR: D60409; D60409.  
DR InterPro: IPR003580; Protachykinin.  
DR InterPro: IPR002040; Tachykinin.  
DR Pfam: PF02202; Tachykinin; 1.  
DR SMART: SM00203; TK; 1.  
DR PROSITE: PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation.  
FT MOD\_RES 11 11 PYRROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 11 11 AMIDATION.  
SQ SEQUENCE 11 AA; 1268 MW; 3DBA7C37C9CB1457 CRC64;  
  
Query Match 59.0%; Score 36; DB 1; Length 11;  
Best Local Similarity 54.5%; Pred. No. 0.6;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKQQOFFGLM 11  
: | : | | |  
DB 1 QPHNPFVGLM 11  
  
RESULT 22  
TKN2\_KASMA STANDARD; PRT; 12 AA.  
AC P08614;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE HYLAMBATIN.  
OS Kassina maculata (African rhacophorid frog) (Hylambates maculatus).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Hyperoliidae;  
OC Kassina.  
OX NCBI\_TaxID=8414;  
RN [1]  
RP SEQUENCE.  
RA yasuhara T., Nakajima T., Erspamer G.F., Erspamer V.;  
RT "New tachykinins, Glu2, Pro5-kassinin (hylambates-kassinin) and  
hylambatin, in the skin of the African rhacophorid frog Hylambates  
maculatus.";  
RL Biomed. Res. 2:613-617(1981).  
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
MUSCLES.  
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR: S07436; S07436.  
DR InterPro: IPR003580; Protachykinin.  
DR InterPro: IPR002040; Tachykinin.

DR Pfam: PF02202; Tachykinin; 1.  
DR SMART: SM00203; TK; 1.  
DR PROSITE: PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.  
FT MOD\_RES 12 12 AMIDATION.  
SQ SEQUENCE 12 AA; 1441 MW; 3287CD2F0DD40AB7 CRC64;  
  
Query Match 59.0%; Score 36; DB 1; Length 12;  
Best Local Similarity 50.0%; Pred. No. 0.66;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 PKPQQOFFGLM 11  
: | : | | |  
DB 3 PDPDFYGM 12  
  
RESULT 23  
TKN\_KASSE STANDARD; PRT; 12 AA.  
AC P08611;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-AUG-1988 (Rel. 08, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE KASSININ.  
OS Kassina senegalensis (Senegal running frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidea; Hyperoliidae;  
OC Kassina.  
OX NCBI\_TaxID=8415;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=77246385; PubMed=891753;  
RA Anastasi A., Montecucci P.C., Erspamer V., Visser J.;  
RT "Amino acid composition and sequence of kassinin, a tachykinin  
dodecapeptide from the skin of the African frog Kassina  
senegalensis.";  
RL Experientia 33:857-858(1977).  
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
MUSCLES.  
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR: S07206; S07206.  
DR InterPro: IPR002040; Tachykinin.  
DR PROSITE: PS00267; TACHYKININ; 1.  
KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.  
FT MOD\_RES 12 12 AMIDATION.  
SQ SEQUENCE 12 AA; 1336 MW; 91757AB89DD6DAB5 CRC64;  
  
Query Match 59.0%; Score 36; DB 1; Length 12;  
Best Local Similarity 70.0%; Pred. No. 0.66;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PKPQQOFFGLM 11  
: | : | | |  
DB 3 PKSDQFVGLM 12  
  
RESULT 24  
SERO\_GALME STANDARD; PRT; 167 AA.  
AC O76192;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE SEROIN PRECURSOR (SILK 23 KDA GLYCOPROTEIN).  
OS Galleria mellonella (Wax moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Pyraloidea; Pyralidae; Galleriinae; Galleria.  
OX NCBI\_TaxID=7137;

RN SEQUENCE FROM N.A., AND SEQUENCE OF 18-31.  
RP TISSUE=silk gland;  
RC MEDLINE=98288272; PubMed=9624126;  
RX Zurovec M., Yang C., Kodrik D., Sehna F.;  
RT "Identification of a novel type of silk protein and regulation of its  
RL expression";  
RL J. Biol. Chem. 273:15428-15428(1998).  
CC -1- SUBCELLULAR LOCATION: SECRETED.  
CC -1- TISSUE SPECIFICITY: PRODUCED BY BOTH THE POSTERIOR (PSG) AND  
CC MIDDLE (MSG) SECTIONS OF SILK GLANDS.  
CC -1- DEVELOPMENTAL STAGE: SEROIN mRNA IS HIGH IN THE SILK GLANDS OF  
CC FEEDING LARVAE, DECLINES AT ECDYSIS, REACHES A MAXIMUM DURING  
CC COCOON SPINNING, AND THEREAFTER RAPIDLY DROPS TO AN UNDETECTABLE  
CC LEVEL.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF009828; AAC25171.1; -  
KW Silk; Glycoprotein; Signal; Repeat.  
FT SIGNAL 1 17  
FT CHAIN 18 167 SEROIN.  
FT REPEAT 38 46 1-1.  
FT REPEAT 56 64 1-2.  
FT REPEAT 76 78 2-1.  
FT REPEAT 79 81 2-2.  
FT REPEAT 82 84 2-3.  
FT CARBOHYD 26 26 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CARBOHYD 146 146 N-LINKED (GLCNAC...) (POTENTIAL).  
SQ SEQUENCE 167 AA; 18088 MW; 27AGABE862774EB9 CRC64;

Query Match 59.0%; Score 36; DB 1; Length 167;  
Best Local Similarity 75.0%; Pred. No. 9.2;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFF 8  
: ||| |||  
Db 103 KPKPGQFF 110

RESULT 25  
ISPE\_HAEIN STANDARD; PRT; 313 AA.  
AC P45271;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DE 4-DIPHOSPHOCYTIDYL-2-C-METHYL-D-ERYTHRITOL KINASE (EC 2.7.1.-) (CMK)  
DE (4-(CYTIDINE-5'-DIPHOSPHO)-2-C-METHYL-D-ERYTHRITOL KINASE).  
GN ISPE OR H11608.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA Mckenney K., Sutton G., Fitzhugh W., Field C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus  
RL influenzae Rd.";  
RL Science 269:496-512(1995).  
CC -1- FUNCTION: CATALYZES THE PHOSPHORYLATION OF THE POSITION 2 HYDROXY  
CC GROUP OF 4-DIPHOSPHOCYTIDYL-2-C-METHYL-D-ERYTHRITOL (BY  
CC SIMILARITY).  
CC -1- PATHWAY: DEOXYXYLOSE-5-PHOSPHATE PATHWAY (DXP) OF ISOPRENOID  
CC BIOSYNTHESIS; FOURTH STEP.  
CC -1- SIMILARITY: BELONGS TO THE ISPE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U32834; AAC23252.1; -  
KW TIGR; H11608; -  
KW Transferase; Kinase; Isoprene biosynthesis; ATP-binding;  
KW Complete proteome.  
SQ NP\_BIND 113 123 ATP (POTENTIAL).  
SQ SEQUENCE 313 AA; 34657 MW; 7A84BAACA196821B CRC64;

Query Match 59.0%; Score 36; DB 1; Length 313;  
Best Local Similarity 54.5%; Pred. No. 17;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
: ||| ||| :  
Db 273 RQKPEAFFGV 283

RESULT 26  
EF2\_ARCFU STANDARD; PRT; 728 AA.  
AC Q28385;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ELONGATION FACTOR 2 (EF-2).  
GN FUSA OR FUS OR AF1894.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
OC Archaeoglobus.  
OX NCBI\_TaxID=2234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE=98049343; PubMed=9389475;  
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,  
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
RA Kirkness E.F., Dougherty B.A., Mckenney K., Adams M.D., Loftus B.,  
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,  
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Uterback T.,  
RA Cotton M.D., Spriggs T., Artlich P., Kaine B.P., Sykes S.M.,  
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;  
RT "The complete genome sequence of the hyperthermophilic, sulphate-  
RT reducing archaeon Archaeoglobus fulgidus.";  
RL Nature 390:364-370(1997).  
CC -1- FUNCTION: THIS PROTEIN PROMOTES THE GTP-DEPENDENT TRANSLLOCATION  
CC OF THE NASCENT PROTEIN CHAIN FROM THE A-SITE TO THE P-SITE OF THE  
CC RIBOSOME.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.  
CC EF-G/EF-2 SUBFAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AE000972; AAB89360.1; -  
CC HSP; P13551; IDAR.  
CC TIGR; AF1894; -  
CC InterPro; IPR000640; EFG\_C.  
CC InterPro; IPR000795; GTP\_EFTU.  
CC Pfam; PF00679; EFG\_C; 1.  
CC Pfam; PF00009; GTP\_EFTU; 1.  
CC PROSITE; PS00301; EFACOR\_GTP; 1.  
CC Elongation factor; Protein biosynthesis; GTP-binding;  
CC Complete proteome.  
CC NP\_BIND 27 34 GTP (BY SIMILARITY).  
CC NP\_BIND 93 97 GTP (BY SIMILARITY).  
CC NP\_BIND 147 150 GTP (BY SIMILARITY).  
CC MOD\_RES 594 594 DIPHTAMIDE (BY SIMILARITY).  
CC SEQUENCE 728 AA; 81135 MW; 62BA963D6571AE9C CRC64;  
-----  
Query Match 59.08; Score 36; DB 1; Length 728;  
Best Local Similarity 66.78; Pred. No. 40;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
-----  
Qy 2 RPKQPFGL 10  
Db 720 RPKDFVGL 728  
|||:|  
-----  
RESULT 27  
SYK\_VIBCH STANDARD; PRT; 512 AA.  
ID SYK\_VIBCH STANDARD; PRT; 512 AA.  
AC Q9KU60;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE) (LYSRS).  
GN LYS OR VC0664.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=686;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Basso S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
RT cholerae";  
RL Nature 406:477-483(2000).  
CC -1- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) = AMP +  
CC PYROPHOSPHATE + L-LYSYL-TRNA(LYS).  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

-----  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AE004152; AAF93829.1; -  
CC TIGR; VC0664; -  
CC InterPro; IPR002106; AA-trna\_ligase\_II.  
CC InterPro; IPR002309; trna-synt\_2.  
CC InterPro; IPR002312; trna-synt\_2.  
CC InterPro; IPR002313; trna-synt\_lys\_2.  
CC Pfam; PF00152; trna-synt\_2; 1.  
CC Pfam; PF01336; trna-anti; 1.  
CC PROSITE; PS00179; AA-trna\_ligase\_II\_1; 1.  
CC PROSITE; PS00339; AA-trna\_ligase\_II\_2; 1.  
CC Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding;  
CC Complete proteome.  
CC SEQUENCE 512 AA; 58336 MW; 3E0BAC911129554C CRC64;  
-----  
Query Match 57.48; Score 35; DB 1; Length 512;  
Best Local Similarity 60.08; Pred. No. 43;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
-----  
Qy 1 RPKQPFGL 10  
Db 152 RPLPEKPHGL 161  
|||:|  
-----  
RESULT 28  
CP3R\_ONCMY STANDARD; PRT; 518 AA.  
ID CP3R\_ONCMY STANDARD; PRT; 518 AA.  
AC O42563;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME P450 3A27 (EC 1.14.14.1) (CYP11A27).  
GN CYP3A27.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
OX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SHASTA; TISSUE=Liver;  
RX MEDLINE=99045386; PubMed=9826429;  
RA Lee S.-J., Wang-Buhler J.-L., Cok I., Yu T.-S., Yang Y.H.,  
RA Miranda C.L., Lech J., Buhler D.R.;  
RT "Cloning, sequencing, and tissue expression of CYP3A27, a new member  
RT of the CYP3A subfamily from embryonic and adult rainbow trout  
RT livers";  
RL Arch. Biochem. Biophys. 360:53-61(1998).  
CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN  
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY  
CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY  
CC ACIDS, AND XENOBIOTICS.  
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -1- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER  
CC TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,  
CC AND CARCINOGENS.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; U96077; AAB82422.1; -  
CC InterPro; IPR001128; Cyt\_P450.

DR Pfam; PF00067; p450; 1.  
DR PRINTS; PR00359; BP450.  
DR PRINTS; PR00385; P450.  
DR PRINTS; PR00463; EP450I.  
DR PRINTS; PR00464; EP450II.  
DR PRINTS; PR00465; EP450IV.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT BINDING 447 447 HEME (BY SIMILARITY).  
SQ SEQUENCE 518 AA; 59210 MW; 9B93AA12E617D0DF CRC64;

Query Match 57.4%; Score 35; DB 1; Length 518;  
Best Local Similarity 60.0%; Pred. No. 43;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKQOFFGLM 11  
DB 42 PKPLPYFGTM 51

RESULT 29  
RELA\_MYXXA STANDARD; PRT; 757 AA.  
ID RELA\_MYXXA  
AC Q52177;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DE GTP PYROPHOSPHOKINASE (EC 2.7.6.5) (ATP:GTP 3'-PYROPHOSPHOTRANSFERASE)  
DE (PPGPP SYNTHETASE I) ((PPGPP SYNTHETASE)).  
GN RELA.  
OS Myxococcus xanthus  
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;  
OC Myxococcales; Cystobacterineae; Myxococcaceae; Myxococcus.  
OX NCBI\_TaxID=34;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DK 101;  
RA Harris B.Z., Kaiser D., Singer M.H.;  
RT "The guanosine nucleotide (ppGpp) initiates development and A-factor  
production in Myxococcus xanthus";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: IN EUBACTERIA PPGBP (GUANOSINE 3'-DIPHOSPHATE 5'-  
DIPHOSPHATE) IS A MEDIATOR OF THE STRINGENT RESPONSE THAT  
COORDINATES A VARIETY OF CELLULAR ACTIVITIES IN RESPONSE TO  
CHANGES IN NUTRITIONAL ABUNDANCE. THIS ENZYME CATALYSES THE  
FORMATION OF PPGBP WHICH IS THEN HYDROLYSED TO FORM PPGBP (BY  
SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + GTP = AMP + GUANOSINE 3'-DIPHOSPHATE  
5'-TRIPHOSPHATE.  
CC -1- PATHWAY: FIRST STEP IN THE METABOLISM OF PPGBP.  
CC -1- SIMILARITY: BELONGS TO THE RELA / SPOT FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; AF025847; AAB97677.1; .  
DR InterPro; IPR002912; ACT.  
DR InterPro; IPR003607; HDC.  
DR Pfam; PF01842; ACT; 1.  
DR SMART; SM00471; HDC; 1.  
KW Transferase; Kinase.  
SQ SEQUENCE 757 AA; 84978 MW; D6CC1000A5F72A7B CRC64;

Query Match 57.4%; Score 35; DB 1; Length 757;  
Best Local Similarity 75.0%; Pred. No. 64;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFFGL 10  
DB 212 KPOEFFAL 219

RESULT 30  
DP2L\_METH STANDARD; PRT; 1092 AA.  
ID DP2L\_METH  
AC Q27579;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DNA POLYMERASE II LARGE SUBUNIT (EC 2.7.7.7) (POL II).  
GN POLC OR MTH1536.  
OS Methanobacterium thermoautotrophicum.  
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
OC Methanothermobacter.  
OX NCBI\_TaxID=145262;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DELTA H;  
RX MEDLINE=98037514; PubMed=9371463;  
RA Smith D.R., Doucette-Stamm L.A., DeLoughery C., Lee H.-M., DuBois J.,  
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Ciu D.,  
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;  
RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
deltaH: functional analysis and comparative genomics";  
RT J. Bacteriol. 179:7135-7155(1997).  
CC -1- FUNCTION: POSSESSES TWO ACTIVITIES: A DNA SYNTHESIS (POLYMERASE)  
AND AN EXONUCLEOTIC ACTIVITY THAT DEGRADES SINGLE STRANDED DNA  
IN THE 3' TO 5' DIRECTION. HAS A TEMPLATE-PRIMER PREFERENCE WHICH  
IS CHARACTERISTIC OF A REPLICATIVE DNA POLYMERASE (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: N DEOXYNUCLEOSIDE TRIPHOSPHATE " N  
PYROPHOSPHATE + DNA(N).  
CC -1- CATALYTIC ACTIVITY: DEGRADATION OF SINGLE-STRANDED DNA. IT ACTS  
PROGRESSIVELY IN A 3' TO 5' DIRECTION, RELEASING 5'-  
PHOSPHONONUCLEOTIDES.  
CC -1- SUBUNIT: HETERODIMER OF A LARGE SUBUNIT AND A SMALL SUBUNIT (BY  
SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE ARCHAEAL DNA POLYMERASE II FAMILY.

-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; A5000913; AAB86010.1; .  
DR Transferase; DNA-directed DNA polymerase; DNA replication; Hydrolase;  
KW Nuclease; Exonuclease; DNA-binding; Multifunctional enzyme;  
KW Complete proteome.  
SQ SEQUENCE 1092 AA; 123058 MW; AA6970F7A6F42DFF CRC64;

Query Match 57.4%; Score 35; DB 1; Length 1092;  
Best Local Similarity 66.7%; Pred. No. 92;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFFGLM 11  
DB 932 KPEQYTGML 940

RESULT 31  
Y130\_MYCPN

ID Y130\_MYCPN STANDARD; PRT; 493 AA.  
AC P75506;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL PROTEIN MG130 HOMOLOG (A65\_ORF493).  
GN MPN269 OR MP564.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 29342 / M129;  
RX MEDLINE=97105885; PubMed=8948633;  
RA Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,  
RA Herrmann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
pneumoniae.";  
RL Nucleic Acids Res. 24:4420-4449(1996).  
CC -!- SIMILARITY: BELONGS TO THE UPF0144 FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 HD DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 HD DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AE000056; AAB96212.1; -  
DR InterPro: IPR002819; HD.  
DR InterPro: IPR003607; HD.  
DR Pfam: PF01966; HD; 1.  
DR SMART: SM00471; Hdc; 1.  
KW Hypothetical protein; Transmembrane; RNA-binding; Complete proteome.  
FT TRANSMEM 19 39 POTENTIAL.  
FT DOMAIN 172 241 KH.  
FT DOMAIN 300 392 HD.  
SQ SEQUENCE 493 AA; 56527 MW; E2D6DE7C8E2FE054 CRC64;  
  
Query Match 55.7%; Score 34; DB 1; Length 493;  
Best Local Similarity 60.0%; Pred. No. 53;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 PKPQOFFGLM 11  
||:||||:  
DB 281 PKLHKFFGLL 290  
  
RESULT 32  
C6B1\_PAPPO STANDARD; PRT; 498 AA.  
AC Q04552; Q04553; Q27878;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME P450 6B1 (EC 1.14.14.1) (CYPVIB1) (CYP6B1V1/CYP6B1V2/  
DE CYP6B1V3).  
GN CYP6B1.  
OS Papilio polyxenes (Black swallowtail butterfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Papilionoidea; Papilionidae; Papilioninae; Papilio.  
OX NCBI\_TaxID=7146;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-25 AND 401-406.  
RC TISSUE=Midgut;  
RX MEDLINE=93066355; PubMed=1279697;  
RA Cohen M.B., Schuler M.A., Berenbaum M.R.;

"A host-inducible cytochrome P-450 from a host-specific caterpillar:  
molecular cloning and evolution.";  
Proc. Natl. Acad. Sci. U.S.A. 89:10920-10924(1992).  
[2]  
RP SEQUENCE FROM N.A. (CYP6B1V3).  
RC TISSUE=Midgut;  
RX MEDLINE=94344788; PubMed=8065937;  
RA Prapalpong H.H., Berenbaum M.M., Schuler M.M.;  
RT "Transcriptional regulation of the Papilio polyxenes CYP6B1 gene.";  
RL Nucleic Acids Res. 22:3210-3217(1994).  
CC -!- FUNCTION: ENABLES THE INSECT TO FEED ON FURANOCUMARIN-PRODUCING  
CC PLANTS AND EVOLVED AS AN ADAPTATION FOR DETOXIFICATION OF  
CC XANTHOTOXIN AND OTHER FURANOCUMARINS.  
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -!- TISSUE SPECIFICITY: MIDGUT MICROSOOME.  
CC -!- INDUCTION: BY XANTHOTOXIN, A SECONDARY METABOLITE ABUNDANT IN THE  
CC HOST PLANTS OF THIS SPECIALIZED HERBIVORE.  
CC -!- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF 6B1-1, 6B1-2 SEEMS  
CC TO DIFFER IN 9 POSITIONS AND IS PROBABLY AN ALLELE.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M80828; AAA29789.1; -  
DR EMBL: M83117; AAA29790.1; -  
DR EMBL: 229624; CA82732.1; -  
DR EMBL: U05037; AAA16154.1; -  
DR PIR: A46367; A46367.  
DR InterPro: IPR001128; Cyt\_P450.  
DR Pfam: PF00067; p450; 1.  
DR PRINTS: PR00385; P450.  
DR PRINTS: PR00464; EP450II.  
DR PROSITE: PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum; Polymorphism.  
FT BINDING 443 443 HEME (BY SIMILARITY).  
FT VARIANT 24 24 N -> D (IN 6B1-2 AND 6B1-3).  
FT VARIANT 155 156 NS -> KC (IN 6B1-2 AND 6B1-3).  
FT VARIANT 243 243 K -> R (IN 6B1-2).  
FT VARIANT 285 285 A -> S (IN 6B1-2).  
FT VARIANT 293 293 I -> V (IN 6B1-2).  
FT VARIANT 458 458 M -> V (IN 6B1-2).  
FT VARIANT 475 475 P -> E (IN 6B1-2).  
FT VARIANT 495 495 L -> I (IN 6B1-2).  
SQ SEQUENCE 498 AA; 57483 MW; 043A849CA0990153 CRC64;  
  
Query Match 55.7%; Score 34; DB 1; Length 498;  
Best Local Similarity 75.0%; Pred. No. 64;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 PKPQOFFG 9  
||| |||  
DB 34 PKPVPFFG 41  
  
RESULT 33  
CP3A\_MESAU  
ID CP3A\_MESAU STANDARD; PRT; 503 AA.  
AC Q64148;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME P450 3A10 (EC 1.14.14.1) (CYP3A10) (6 BETA-HYDROXYLASE).  
GN CYP3A10.

OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus  
OX NCBI\_TaxID=10036;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95379809; PubMed=7651384;  
RA Subramanian A., Teixeira J., Wang J., Gil G.;  
"A STAT factor mediates the sexually dimorphic regulation of hepatic  
RT cytochrome P450 3A10/lithocholic acid 6 beta-hydroxylase gene  
RT expression by growth hormone.";  
Mol. Cell. Biol. 15:4672-4682(1995).  
CC -1- FUNCTION: CATALYZES THE 6 BETA-HYDROXYLATION OF LITHOCHOLIC ACID  
CC AND STEROID HORMONES.  
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED ONLY IN MALE HAMSTERS.  
CC -1- INDUCTION: P450 CAN BE INDUCED TO HIGH LEVELS IN LIVER AND OTHER  
CC TISSUES BY VARIOUS FOREIGN COMPOUNDS, INCLUDING DRUGS, PESTICIDES,  
CC AND CARCINOGENS.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M73992; ; NOT\_ANNOTATED\_CDS.  
CC EMBL; M73917; AAB35091.1; -;  
CC InterPro; IPR001128; Cyt\_P450.  
CC Pfam; PF00067; p450; 1.  
CC PRINTS; PR00359; BP450.  
CC PRINTS; PR00385; P450.  
CC PRINTS; PR00463; EP450I.  
CC PRINTS; PR00464; EP450II.  
CC PRINTS; PR00465; EP450IV.  
CC PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum.  
FT BINDING 442 442 HEME (BY SIMILARITY).  
FT SEQUENCE 503 AA; 57693 MW; D4D24FEE87FD7F51 CRC64;

Query Match 55.7%; Score 34; DB 1; Length 503;  
Best Local Similarity 75.0%; Pred. No. 65;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQQFFG 9  
||| |||  
Db 41 PKPLPFFG 48

RESULT 34  
CP31\_RAT  
ID CP31\_RAT STANDARD; PRT; 504 AA.  
AC P04800; Q64580;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 13-AUG-1987 (Rel. 05, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME P450 3A1 (EC 1.14.14.1) (CYP11A1) (P450-PCN1).  
GN CYP3A1 OR CYP3A-1.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=101116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85207783; PubMed=3838989;  
RA Gonzalez F.J., Nebert D.W., Hardwick J.P., Kasper C.B.;  
"Complete cDNA and protein sequence of a pregnenolone 16 alpha-  
RT carbonitrile-induced cytochrome P-450. A representative of a new gene  
RT family.";  
J. Biol. Chem. 260:7435-7441(1985).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=WISTAR; TISSUE=Liver;  
MEDLINE=92117688; PubMed=1731631;  
RA Ribeiro V., Lechner M.C.;  
"Cloning and characterization of a novel CYP3A1 allelic variant:  
RT analysis of CYP3A1 and CYP3A2 sex-hormone-dependent expression reveals  
RT that the CYP3A2 gene is regulated by testosterone.";  
Arch. Biochem. Biophys. 293:147-152(1992).  
RN [3]  
RP SEQUENCE OF 1-24 FROM N.A.  
RX MEDLINE=92196074; PubMed=1372436;  
RA Burger H.J., Schuetz J.D., Schuetz E.G., Guzelian P.S.;  
"Paradoxical transcriptional activation of rat liver cytochrome P-450  
RT 3A1 by dexamethasone and the antiglucoctericoid pregnenolone 16  
RT alpha- carbonitrile: analysis by transient transfection into primary  
RT monolayer cultures of adult rat hepatocytes.";  
Proc. Natl. Acad. Sci. U.S.A. 89:2145-2149(1992).  
CC -1- FUNCTION: CYTOCHROMES P450 ARE A GROUP OF HEME-THIOLATE  
CC MONOOXYGENASES. IN LIVER MICROSOMES, THIS ENZYME IS INVOLVED IN AN  
CC NADPH-DEPENDENT ELECTRON TRANSPORT PATHWAY. IT OXIDIZES A VARIETY  
CC OF STRUCTURALLY UNRELATED COMPOUNDS, INCLUDING STEROIDS, FATTY  
CC ACIDS, AND XENOBIOTICS.  
CC -1- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -1- INDUCTION: BY PREGNENOLONE 16-ALPHA-CARBONITRILE (PNCN).  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M10161; AAA41035.1; -;  
CC EMBL; X64401; CAA45743.1; -;  
CC EMBL; M86850; AAA41780.1; -;  
CC PIR; A22631; A22631.  
CC InterPro; IPR001128; Cyt\_P450.  
CC Pfam; PF00067; p450; 1.  
CC PRINTS; PR00385; P450.  
CC PRINTS; PR00464; BP450II.  
CC PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum; Polymorphism.  
FT BINDING 443 443 HEME.  
FT VARIANT 207 207 T -> A (IN STRAIN WISTAR).  
FT VARIANT 213 213 F -> I (IN STRAIN WISTAR).  
FT VARIANT 232 232 I -> V (IN STRAIN WISTAR).  
FT SEQUENCE 504 AA; 57917 MW; CFF5AC8C37E9CADB CRC64;

Query Match 55.7%; Score 34; DB 1; Length 504;  
Best Local Similarity 75.0%; Pred. No. 65;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQQFFG 9  
||| |||



```
PPOB_SOLTU          STANDARD;          PRT;          588 AA.
ID  PPOB_SOLTU          STANDARD;          PRT;          588 AA.
AC  Q06355;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  01-NOV-1995 (Rel. 32, Last annotation update)
DE  POLYPHENOL OXIDASE B PRECURSOR (EC 1.10.3.1) (PPO) (CATECHOL OXIDASE)
DE  (FRAGMENT).
OS  Solanum tuberosum (Potato).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC  Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX  NCBI_TaxID=4113;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=CV, KATAHDIN; TISSUE=Leaf;
RX  MEDLINE=93144692; PubMed=7678763;
RA  Hunt M.D., Eannetta N.T., Yu H., Newman S.M., Steffens J.C.;
RA  "cDNA cloning and expression of potato polyphenol oxidase.";
RL  Plant Mol. Biol. 21:59-68(1993).
CC  -1- FUNCTION: CATALYZE THE OXIDATION OF MONO- AND O-DIPHENOLS TO O-
CC  DIQUINONES.
CC  -1- CATALYTIC ACTIVITY: 2 CATECHOL + O(2) = 2 1,2-BENZQUINONE +
CC  2 H(2)O.
CC  -1- COFACTOR: BINDS TWO COPPER IONS (BY SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: CHLOROPLAST; WITHIN THE THYLAKOID LUMEN.
CC  -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M95197; AAA02879.1; -.
DR  InterPro; IPR002227; Tyrosinase.
DR  Pfam; PF00264; tyrosinase; 1.
DR  PROSITE; PS00497; TYROSINASE.1; 1.
DR  PROSITE; PS00498; TYROSINASE.2; 1.
KW  Oxidoreductase; Copper; Metal-binding; Chloroplast; Transit peptide;
KW  Multigene family.
FT  NON_TER 1
FT  TRANSIT <1 88 CHLOROPLAST (POTENTIAL).
FT  CHAIN 89 588 POLYPHENOL OXIDASE B.
FT  METAL 198 198 COPPER A (BY SIMILARITY).
FT  METAL 207 207 COPPER A (BY SIMILARITY).
FT  METAL 329 329 COPPER B (BY SIMILARITY).
FT  METAL 333 333 COPPER B (BY SIMILARITY).
FT  METAL 364 364 COPPER B (BY SIMILARITY).
SQ  SEQUENCE 588 AA; 66240 MW; A7E25383273428CC CRC64;

Query Match 55.7%; Score 34; DB 1; Length 588;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQFFG 9
   | | | | |
DB 301 PCPSQFFG 308

RESULT 38
YJB0_YEAST          STANDARD;          PRT;          666 AA.
ID  YJB0_YEAST          STANDARD;          PRT;          666 AA.
AC  P47077;
DT  01-FEB-1996 (Rel. 33, Created)
DT  01-FEB-1996 (Rel. 33, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  HYPOTHEICAL 77.7 KDA PROTEIN IN CCT3-CCT8 INTERGENIC REGION.
GN  YJL010C OR J1357.
OS  Saccharomyces cerevisiae (Baker's yeast).

PPOB_SOLTU          STANDARD;          PRT;          588 AA.
ID  PPOB_SOLTU          STANDARD;          PRT;          588 AA.
AC  Q06355;
DT  01-NOV-1995 (Rel. 32, Created)
DT  01-NOV-1995 (Rel. 32, Last sequence update)
DT  01-NOV-1995 (Rel. 32, Last annotation update)
DE  POLYPHENOL OXIDASE B PRECURSOR (EC 1.10.3.1) (PPO) (CATECHOL OXIDASE)
DE  (FRAGMENT).
OS  Solanum tuberosum (Potato).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC  Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX  NCBI_TaxID=4113;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=CV, KATAHDIN; TISSUE=Leaf;
RX  MEDLINE=93144692; PubMed=7678763;
RA  Hunt M.D., Eannetta N.T., Yu H., Newman S.M., Steffens J.C.;
RA  "cDNA cloning and expression of potato polyphenol oxidase.";
RL  Plant Mol. Biol. 21:59-68(1993).
CC  -1- FUNCTION: CATALYZE THE OXIDATION OF MONO- AND O-DIPHENOLS TO O-
CC  DIQUINONES.
CC  -1- CATALYTIC ACTIVITY: 2 CATECHOL + O(2) = 2 1,2-BENZQUINONE +
CC  2 H(2)O.
CC  -1- COFACTOR: BINDS TWO COPPER IONS (BY SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: CHLOROPLAST; WITHIN THE THYLAKOID LUMEN.
CC  -1- SIMILARITY: BELONGS TO THE TYROSINASE FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; M95197; AAA02879.1; -.
DR  InterPro; IPR002227; Tyrosinase.
DR  Pfam; PF00264; tyrosinase; 1.
DR  PROSITE; PS00497; TYROSINASE.1; 1.
DR  PROSITE; PS00498; TYROSINASE.2; 1.
KW  Oxidoreductase; Copper; Metal-binding; Chloroplast; Transit peptide;
KW  Multigene family.
FT  NON_TER 1
FT  TRANSIT <1 88 CHLOROPLAST (POTENTIAL).
FT  CHAIN 89 588 POLYPHENOL OXIDASE B.
FT  METAL 198 198 COPPER A (BY SIMILARITY).
FT  METAL 207 207 COPPER A (BY SIMILARITY).
FT  METAL 329 329 COPPER B (BY SIMILARITY).
FT  METAL 333 333 COPPER B (BY SIMILARITY).
FT  METAL 364 364 COPPER B (BY SIMILARITY).
SQ  SEQUENCE 588 AA; 66240 MW; A7E25383273428CC CRC64;

Query Match 55.7%; Score 34; DB 1; Length 588;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQFFG 9
   | | | | |
DB 301 PCPSQFFG 308

RESULT 38
YJB0_YEAST          STANDARD;          PRT;          666 AA.
ID  YJB0_YEAST          STANDARD;          PRT;          666 AA.
AC  P47077;
DT  01-FEB-1996 (Rel. 33, Created)
DT  01-FEB-1996 (Rel. 33, Last sequence update)
DT  01-NOV-1997 (Rel. 35, Last annotation update)
DE  HYPOTHEICAL 77.7 KDA PROTEIN IN CCT3-CCT8 INTERGENIC REGION.
GN  YJL010C OR J1357.
OS  Saccharomyces cerevisiae (Baker's yeast).

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
NCBI_TaxID=4932;
[1]
SEQUENCE FROM N.A.
To Van D., Perea J., Jacq C.;
Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: TO S.POMBE SPAC6G9.02C.
-----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; Z49285; CAA89301.1; -.
DR  SGD; S0003547; YJL010C.
DR  InterPro; IPR001313; PUM.
DR  Pfam; PF00806; PUF; 8.
DR  SMART; SM00025; Pumilio; 8.
KW  Hypothetical protein.
SQ  SEQUENCE 666 AA; 77722 MW; F6F8B3CD74DB2AB3 CRC64;

Query Match 55.7%; Score 34; DB 1; Length 666;
Best Local Similarity 55.6%; Pred. No. 86;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPQOFFGLM 11
   ||| |||
DB 48 QPQOFFGLV 56

RESULT 39
Y025_CAEEL          STANDARD;          PRT;          1799 AA.
ID  Y025_CAEEL          STANDARD;          PRT;          1799 AA.
AC  P34675;
DT  01-FEB-1994 (Rel. 28, Created)
DT  01-FEB-1994 (Rel. 28, Last sequence update)
DT  15-JUL-1998 (Rel. 36, Last annotation update)
DE  HYPOTHEICAL 202.6 KDA PROTEIN ZK688.5 IN CHROMOSOME III.
GN  ZK688.5.
OS  Caenorhabditis elegans.
OC  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC  Rhabditidae; Telodierinae; Caenorhabditis.
OX  NCBI_TaxID=6239;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BRISTOL N2;
RX  MEDLINE=94150718; PubMed=7906398;
RA  Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA  Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA  Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA  Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA  Johnston L., Jones M., Kersey J., Kirov J., Laister N.,
RA  Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA  Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkneen R.,
RA  Sims M., Smalton N., Smith A., Smith M., Sonhammer E., Staden R.,
RA  Sulston J., Thierry-Mieg J., Thomas K., Vaughan M., Vaughan K.,
RA  Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA  Wohldman P.;
RA  *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT  elegans.
RL  Nature 368:32-38(1994).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
```



```
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L16621; AAA28231.1; -.
CC PIR; S44920; S44920.
CC WormPep; ZK688.5; CE00463.
CC InterPro; IPR000488; Death.
CC InterPro; IPR000626; Ubiquitin.
CC Pfam; PF00240; ubiquitin; 1.
CC SMART; SM00005; DEATH; 1.
CC SMART; SM00213; UBO; 1.
CC PROSITE; PS00053; UBIQUITIN_2; UNKNOWN_1.
CC KW Ubiquitin-like protein.
FT DOMAIN 21 96 UBIQUITIN-LIKE.
SQ SEQUENCE 1799 AA; 202641 MW; FFF4F79C4B5980B4 CRC64;

Query Match 55.78; Score 34; DB 1; Length 1799;
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQF 7
|||
Db 552 RPTPQQF 558

RESULT 40
YKCO_CAEEL
ID YKCO_CAEEL STANDARD; PRT; 282 AA.
AC P42001;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HYPOTHETICAL 32.9 KDA PROTEIN B0280.10 IN CHROMOSOME III.
GN B0280.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Fulton L., Waterston R.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U10438; AAA19088.1; -.
CC WormPep; B0280.10; CE00817.
CC KW Hypothetical protein.
SQ SEQUENCE 282 AA; 32883 MW; ED3EE1E98889CED7 CRC64;

Query Match 54.1%; Score 33; DB 1; Length 282;
Best Local Similarity 62.5%; Pred. No. 56;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQQFFG 9
| | | | |
Db 15 PTPQNFYG 22

RESULT 41
CAG5_CHICK
ID CAG5_CHICK STANDARD; PRT; 404 AA.
AC Q92184;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
```

```
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE ALPHA-N-ACETYLGALACTOSAMINIDE ALPHA-2,6-SIALYLTRANSFERASE
DE (EC 2.4.99.-) (ST6GALNACII).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Testis;
RL MEDLINE=94308168; PubMed=8034663;
RA Kurosawa N., Kojima N., Inoue M., Hamamoto T., Tsuji S.;
RT "Cloning and expression of Gal beta 1,3galnac-specific Galnac alpha
RT 2,6-sialyltransferase";
RL J. Biol. Chem. 269:19048-19053(1994).
CC -!- CATALYTIC ACTIVITY: CMP-N-ACETYLNEURAMINATE + GLYCANO-BETA-D-
CC GALACTOSYL-1,3-(N-ACETYL-D-GALACTOSAMINYL)-GLYCOPROTEIN = CMP
CC + GLYCANO-BETA-D-GALACTOSYL-(2,6-ALPHA-N-ACETYLNEURAMINYL)-
CC (N-ACETYL-D-GALACTOSAMINYL)-GLYCOPROTEIN.
CC -!- PATHWAY: GLYCOSYLATION.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI (POTENTIAL).
CC -!- TISSUE SPECIFICITY: HEART, KIDNEY, TESTES, BRAIN, LIVER AND LUNG.
CC -!- DEVELOPMENTAL STAGE: ABUNDANTLY EXPRESSED AT ALL EMBRYONIC STAGES
CC BUT NOT PRESENT IN ADULT TISSUES.
CC -!- SIMILARITY: BELONGS TO THE VERTEBRATE SIALYLTRANSFERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X77775; CAA54813.1; -.
CC InterPro; IPR001675; Glyco.transf_29.
CC Pfam; PF00777; Glyco.transf_29; 1.
CC KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
FT SIGNAL-ANCHOR; Golgi stack.
FT DOMAIN 1 8
FT TRANSMEM 9 25 CYTOPLASMIC (POTENTIAL).
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
(POTENTIAL).
FT DOMAIN 26 404 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 191 191 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 404 AA; 45826 MW; DCC177AA01ABB60A CRC64;

Query Match 54.1%; Score 33; DB 1; Length 404;
Best Local Similarity 62.5%; Pred. No. 80;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 KPQQFFGL 10
| | | | |
Db 288 EPQKYFGL 295

RESULT 42
CRF1_CHICK
ID CRF1_CHICK STANDARD; PRT; 420 AA.
AC O80812;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CORTICOTROPIN-RELEASING FACTOR RECEPTOR 1 PRECURSOR (CRF-R) (CRF1)
DE (CORTICOTROPIN-RELEASING HORMONE RECEPTOR 1) (CRH-R 1).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE-96107136; PubMed-8536612;
RA Yu J., Xie L.Y., Abou-Samra A.-B.;
RT "Molecular cloning of a type A chicken corticotropin-releasing factor
RT receptor with high affinity for urotensin I.";
RL Endocrinology 137:192-197(1996).
CC -!- FUNCTION: THIS IS A RECEPTOR FOR CORTICOTROPIN RELEASING FACTOR.
CC SHOWS HIGH-AFFINITY BINDING FOR UROTENSIN I. THE ACTIVITY OF THIS
CC RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE ADENYL
CC CYCLASE.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.
CC -!- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L41563; AAA96656.1;
CC GCRDB; GCR_1180;
CC InterPro; IPR000832; GPCR_secretin.
CC Pfam; PF00002; 7tm_2; 1.
CC PRINTS; PR00249; GPCRSECRETIN.
CC PRINTS; PR01279; CRFRECEPTOR.
CC PRINTS; PR01280; CRFRECEPTOR1.
CC SMART; SM00008; Hormr; 1.
CC PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; 1.
CC PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; 1.
CC PROSITE; PS50227; G_PROTEIN_RECEP_F2_3; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 28
FT CHAIN 29 420
FT FT 29 420
FT DOMAIN 29 126
FT TRANSMEM 127 147
FT DOMAIN 148 156
FT TRANSMEM 157 176
FT DOMAIN 177 194
FT TRANSMEM 195 218
FT DOMAIN 219 232
FT TRANSMEM 233 254
FT DOMAIN 255 273
FT TRANSMEM 274 296
FT DOMAIN 297 319
FT TRANSMEM 320 339
FT DOMAIN 340 354
FT TRANSMEM 355 374
FT DOMAIN 375 420
FT CARBOHYD 43 43
FT CARBOHYD 50 50
FT CARBOHYD 83 83
FT CARBOHYD 95 95
FT CARBOHYD 103 103
FT SEQUENCE 420 AA; 48600 MW; 8C5C992925F62316 CRC64;

Query Match 54.1%; Score 33; DB 1; Length 420;
Best Local Similarity 50.0%; Pred. No. 83;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQOFFGL 10
|| | | |
DB 71 RPKPEYFGV 80

RESULT 43
SYN_STAAU STANDARD; PRT; 495 AA.
AC Q53638;
```

```
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE) (LYSRS).
GN LYSS.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 12600;
RA Green C.J., Vold B.S.;
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) = AMP +
CC PYROPHOSPHATE + L-LYSYL-TRNA(LYS).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L36472; AAA53114.1;
CC HSSP; PI4825; LLYL.
CC InterPro; IPR002106; AA_trna_ligase-II.
CC InterPro; IPR002309; trna-synt_2.
CC InterPro; IPR002313; trna-synt_lys_2.
CC Pfam; PF00152; trna-synt_2; 1.
CC Pfam; PF01336; trna-anti; 1.
CC PRINTS; PR00982; TRNASYNTHLYS.
CC PROSITE; PS00179; AA_TRNA_LIGASE-II_1; 1.
CC PROSITE; PS00339; AA_TRNA_LIGASE-II_2; 1.
CC Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.
KW SEQUENCE 495 AA; 56889 MW; E5DC43BECB0E9D98 CRC64;
SQ

Query Match 54.1%; Score 33; DB 1; Length 495;
Best Local Similarity 60.0%; Pred. No. 98;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQOFFGL 10
|| | | |
DB 143 RPLPDRFHL 152

RESULT 44
C6B3_PAPPO STANDARD; PRT; 498 AA.
ID C6B3_PAPPO
AC Q27756; Q95038;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CYTOCHROME P450 6B3 (EC 1.14.14.1) (CYP6B3V1/CYP6B3V2).
OS CYP6B3.
OS Papilio polyxenes (Black swallowtail butterfly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;
OC Papilionoidea; Papilionidae; Papilioninae; Papilio.
OX NCBI_TaxID=7146;
RN [1]
RP SEQUENCE FROM N.A. (CYP6B3V1).
RX MEDLINE-96111294; PubMed-8589841;
RA Hung C.F., Harrison T.L., Berenbaum M.R., Schuler M.A.;
RT "CYP6B3: a second furanocoumarin-inducible cytochrome P450 expressed
RT in Papilio polyxenes.";
RL Insect Mol. Biol. 4:149-160(1995).
RN [2]
```

RP SEQUENCE FROM N.A. (CYP6B3V2).  
RX MEDLINE=97057218; PubMed=8901557;  
RA Hung C.F., Holznacher R., Connolly E., Berenbaum M.R., Schuler M.A.;  
RT "Conserved promoter elements in the CYP6B gene family suggest common  
ancestry for cytochrome P450 monooxygenases mediating furanocoumarin  
detoxification.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:12200-12205(1996).  
CC -!- FUNCTION: ENABLES THE INSECT TO FEED ON FURANOCOUMARIN-PRODUCING  
PLANTS AND EVOLVED AS AN ADAPTATION FOR DETOXIFICATION OF  
XANTHOTOXIN AND OTHER FURANOCOUMARINS.  
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +  
OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM.  
CC -!- INDUCTION: BY XANTHOTOXIN AND BERGAPTEN (LINEAR FURANOCOUMARINS)  
AS WELL AS BY ANGELICIN AND SPHONDIN (ANGULAR FURANOCOUMARINS).  
CC -!- POLYMORPHISM: THE SEQUENCE SHOWN IS THAT OF 6B3-1, 6B3-2 SEEMS  
TO DIFFER IN 17 POSITIONS AND IS PROBABLY AN ALLELE.  
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U25819; AAA96255.1; -;  
DR EMBL; U65488; AAB06741.1; -;  
DR InterPro; IPR001128; Cyt\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00385; P450.  
DR PRINTS; PR00464; EP450II.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;  
KW Microsome; Endoplasmic reticulum; Polymorphism.  
FT BINDING 443 443 HEME (BY SIMILARITY).  
FT VARIANT 47 47 A -> V (IN 6B3-2).  
FT VARIANT 52 52 P -> H (IN 6B3-2).  
FT VARIANT 82 82 L -> I (IN 6B3-2).  
FT VARIANT 92 93 PT -> LI (IN 6B3-2).  
FT VARIANT 98 98 P -> S (IN 6B3-2).  
FT VARIANT 108 108 L -> I (IN 6B3-2).  
FT VARIANT 289 289 T -> I (IN 6B3-2).  
FT VARIANT 350 350 G -> S (IN 6B3-2).  
FT VARIANT 354 357 FLGR -> YLSK (IN 6B3-2).  
FT VARIANT 395 397 IIV -> VII (IN 6B3-2).  
FT VARIANT 401 401 G -> S (IN 6B3-2).  
SQ SEQUENCE 498 AA; 57473 MW; 9BC760ACBEB657BC CRC64;  
  
Query Match 54.1%; Score 33; DB 1; Length 498;  
Best Local Similarity 75.0%; Pred. No. 98;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 PKPQOFFG 9  
DB 34 PRPIFFG 41  
  
RESULT 45  
SYK\_PASMU STANDARD; PRT; 501 AA.  
AC P57822;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE) (LYSRS).  
GN LYSS OR LYSU OR PM0189.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=747;

[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=PM70;  
RX MEDLINE=21145866; PubMed=11248100;  
RA May B.J., Zhang Q., Li L.L., Faustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -!- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) = AMP +  
PYROPHOSPHATE + L-LYSYL-TRNA(LYS).  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE006053; AAK02273.1; -;  
DR InterPro; IPR002106; AA\_trna\_ligase\_II.  
DR InterPro; IPR002309; trna-synt\_2.  
DR InterPro; IPR002313; trna-synt\_lys\_2.  
DR Pfam; PF00152; trna-synt\_2; 1.  
DR Pfam; PF01336; trna-anti; 1.  
DR PROSITE; PS00179; AA\_trna\_ligase\_II\_1; 1.  
DR PROSITE; PS00339; AA\_trna\_ligase\_II\_2; 1.  
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
SQ SEQUENCE 501 AA; 56759 MW; 767E6F21AC85B240 CRC64;  
  
Query Match 54.1%; Score 33; DB 1; Length 501;  
Best Local Similarity 60.0%; Pred. No. 99;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKQOFFGL 10  
DB 146 RPLPKRFHGL 155  
  
RESULT 46  
SYK\_HAEIN  
ID SYK\_HAEIN STANDARD; PRT; 502 AA.  
AC P43825;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE) (LYSRS).  
GN LYSS OR LYSU OR H1211.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Cocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus  
influenzae Rd.";  
RL Science 269:496-512(1995).  
CC -!- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) = AMP +

CC PYROPHOSPHATE + L-LYSYL-TRNA(LYS).  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U32800; AAC22865.1; -.  
DR HSSP; P14825; 1LYL.  
DR TIGR; H11211; -.  
DR InterPro; IPR002106; AA\_trna\_ligase\_II.  
DR InterPro; IPR002309; trna-synt\_2.  
DR InterPro; IPR002313; trna-synt\_lys\_2.  
DR Pfam; PF00152; trna-synt\_2; 1.  
DR Pfam; PF01336; trna-anti; 1.  
DR PRINTS; PR00982; TRNASYNTHLYS.  
DR PROSITE; PS00179; AA\_TRNA\_LIGASE\_II\_1; 1.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; 1.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.  
SQ SEQUENCE 502 AA; 56935 MW; DF281DF073A702B9 CRC64;

Query Match 54.1%; Score 33; DB 1; Length 502;  
Best Local Similarity 60.0%; Pred. No. 99;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQFFGL 10  
|||:|:|  
DB 147 RPLPKPHGL 156

RESULT 47  
SYKL\_ECOLI  
ID SYKL\_ECOLI STANDARD; PRT; 504 AA.  
AC P13030;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE) (LYSRS).  
GN LYSS OR HEC OR ASUD OR B2890.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
OX [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=88289768; PubMed=2456575;  
RA Kawakami K., Joensson Y.H., Bjoerk G.R., Ikeda H., Nakamura Y.;  
RT "Chromosomal location and structure of the operon encoding  
RT peptide-chain-release factor 2 of Escherichia coli.";  
RL Proc. Natl. Acad. Sci. U.S.A. 85:5620-5624(1988).  
RN [2]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-27.  
RX MEDLINE=90221811; PubMed=2183178;  
RA Leveque F., Plateau P., Dessen P., Blanquet S.;  
RT "Homology of lysyl and lysu, the two Escherichia coli genes encoding  
RT distinct lysyl-tRNA synthetase species.";  
RL Nucleic Acids Res. 18:305-312(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN [4]  
RP STRUCTURE BY NMR.  
RX MEDLINE=96028214; PubMed=7473706;  
RA Commans S., Plateau P., Blanquet S., Dardel F.;  
RT "Solution structure of the anticodon-binding domain of Escherichia  
RT coli lysyl-tRNA synthetase and studies of its interaction with  
RT tRNA(Lys).";  
RL J. Mol. Biol. 253:100-113(1995).  
CC -1- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) -> AMP +  
CC -1- PYROPHOSPHATE + L-LYSYL-TRNA(LYS).  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- MISCELLANEOUS: THERE ARE TWO LYSYL-TRNA LIGASES IN E. COLI: LYSS IS  
CC EXPRESSED CONSTITUTIVELY, WHILE LYSU IS HEAT INDUCIBLE.  
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J03795; AAA23959.1; -.  
DR EMBL; U28375; AAA83071.1; -.  
DR EMBL; AE000372; AAC75928.1; -.  
DR PIR; JS0401; SYECKT.  
DR PDB; 1KRS; 15-SEP-95.  
DR PDB; 1KRT; 15-SEP-95.  
DR SWISS-2DPAGE; P13030; COLI.  
DR ECOBASE; D058.5; 6TH EDITION.  
DR Ecogene; EG10552; lyss.  
DR InterPro; IPR002106; AA\_trna\_ligase\_II.  
DR InterPro; IPR002309; trna-synt\_2.  
DR InterPro; IPR002313; trna-synt\_lys\_2.  
DR Pfam; PF00152; trna-synt\_2; 1.  
DR Pfam; PF01336; trna-anti; 1.  
DR PRINTS; PR00982; TRNASYNTHLYS.  
DR PROSITE; PS00179; AA\_TRNA\_LIGASE\_II\_1; 1.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; 1.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Multigene family; 3D-structure; Complete proteome.  
FT INIT\_MET 0  
FT INIT\_MET 0  
SQ SEQUENCE 504 AA; 57472 MW; EE5F3D1FBA63CFEF CRC64;

Query Match 54.1%; Score 33; DB 1; Length 504;  
Best Local Similarity 60.0%; Pred. No. 99;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQFFGL 10  
|||:|:|  
DB 150 RPLPKPHGL 159

## RESULT 48

SYK2\_ECOLI  
ID SYK2\_ECOLI STANDARD; PRT; 504 AA.  
AC P14825;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LYSYL-TRNA SYNTHETASE, HEAT INDUCIBLE (EC 6.1.1.6) (LYSINE--TRNA  
DE LIGASE) (LYSRS).  
GN LYSU OR B4129.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=562;  
OX [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-20.  
RX MEDLINE=90221811; PubMed=2183178;  
RA Leveque F., Plateau P., Dessen P., Blanquet S.;  
RT "Homology of lysS and lysU, the two Escherichia coli genes encoding  
RL distinct lysyl-tRNA synthetase species";  
RN Nucleic Acids Res. 18:305-312(1990).  
RP [2]  
RP REVISION TO 445.  
RA Dessen P.;  
RL Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90264318; PubMed=2188953;  
RA Clark R.L., Neidhardt F.C.;  
RT "Roles of the two lysyl-tRNA synthetases of Escherichia coli:  
RT analysis of nucleotide sequences and mutant behavior";  
RN J. Bacteriol. 172:3237-3243(1990).  
RP [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=95334362; PubMed=7610040;  
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,  
RA Blattner F.R.;  
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the  
RT region from 92.8 through 100 minutes";  
RN Nucleic Acids Res. 23:2105-2119(1995).  
RP [5]  
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).  
RX MEDLINE=95253817; PubMed=7735833;  
RA Onesti S., Miller A.D., Brick P.;  
RT "The crystal structure of the lysyl-tRNA synthetase (LysU) from  
RT Escherichia coli";  
RL Structure 3:163-176(1995).  
CC -!- FUNCTION: ALSO CAN SYNTHESIZE A NUMBER OF ADENYL DINUCLEOTIDES (IN  
CC PARTICULAR APPRA). THESE DINUCLEOTIDES HAVE BEEN PROPOSED TO ACT  
CC AS MODULATORS OF THE HEAT-SHOCK RESPONSE AND STRESS RESPONSE.  
CC -!- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) = AMP +  
CC PYROPHOSPHATE + L-LYSYL-TRNA(LYS).  
CC -!- SUBUNIT: HOMODIMER.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- MISCELLANEOUS: THERE ARE TWO LYSYL-TRNA LIGASES IN E.COLI: LYSS IS  
CC EXPRESSED CONSTITUTIVELY, WHILE LYSU IS HEAT INDUCIBLE.  
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X16542; CAA34542.1; -  
DR EMBL; M30630; AAA24096.1; -  
DR EMBL; U14003; AAA97029.1; -  
DR EMBL; AE000485; AAC77090.1; -  
DR PIR; JS0400; SYECKU  
DR PDB; 1LYL; 15-OCT-95.  
DR ECODBASE; D060.5; 6TH EDITION.  
DR EcoGene; EG10553; lysU.  
DR InterPro; IPR002106; AA\_trna\_ligase\_II.  
DR InterPro; IPR002309; trna-synt\_2.  
DR InterPro; IPR002313; trna-synt\_lys\_2.  
DR Pfam; PF00152; trna-synt\_2; 1.  
DR Pfam; PF01336; trna-anti; 1.  
DR PRINTS; PR00982; TRNASYNTHYS.  
DR PROSITE; PS00179; AA\_TRNA\_LIGASE\_II\_1; 1.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; 1.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Multigene family; 3D-structure; Complete proteome.  
FT INIT\_MET 0  
FT CONFLICT 124 124 MISSING (IN REF. 3).  
FT CONFLICT 235 235 L -> A (IN REF. 3).  
FT

FT CONFLICT 257 261 INRNF -> HVT (IN REF. 3).  
FT CONFLICT 267 268 SV -> R (IN REF. 3).  
FT CONFLICT 350 350 A -> R (IN REF. 3).  
FT CONFLICT 370 370 I -> S (IN REF. 3).  
FT CONFLICT 379 383 AEAHL -> VEGHV (IN REF. 3).  
FT CONFLICT 387 387 T -> S (IN REF. 3).  
SQ SEQUENCE 504 AA; 57695 MW; 48E1B8F5875396E0 CRC64;  
  
Query Match 54.1%; Score 33; DB 1; Length 504;  
Best Local Similarity 60.0%; Pred. No. 99;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKPOQFFGL 10  
Db 150 RPLPKFHGL 159  
|| | : ||  
- - - - -  
RESULT 49  
FRK\_HUMAN  
ID FRK\_HUMAN STANDARD; PRT; 505 AA.  
AC P42685; Q13128;  
DC 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE TYROSINE-PROTEIN KINASE FRK (EC 2.7.1.112) (NUCLEAR TYROSINE PROTEIN  
DE KINASE RAK).  
DE GN FRK.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lymphoid;  
RX MEDLINE=94171047; PubMed=7510261;  
RA Lee J., Wang Z., Luoh S.-M., Wood W.I., Scadden D.T.;  
RT "Cloning of FRK, a novel human intracellular SRC-like tyrosine  
RT kinase-encoding gene";  
RL Gene 138:247-251(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95210168; PubMed=7696183;  
RA Cance W.G., Craven R.J., Bergman M., Xu L.H., Alitalo K., Liu E.T.;  
RT "Rak, a novel nuclear tyrosine kinase expressed in epithelial cells";  
RL Cell Growth Differ. 5:1347-1355(1994).  
RN [3]  
RP PARTIAL SEQUENCE FROM N.A.  
RX MEDLINE=93293373; PubMed=8099900;  
RA Cance W.G., Craven R.J., Weiner T.M., Liu E.T.;  
RT "Novel protein kinases expressed in human breast cancer";  
RL Int. J. Cancer 54:571-577(1993).  
CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE = ADP +  
CC PROTEIN TYROSINE PHOSPHATE.  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).  
CC -!- TISSUE SPECIFICITY: RESTRICTED TO CELLS LINES DERIVED FROM TISSUES  
CC OF LYMPHOID, BRAIN, BREAST, COLON AND BLADDER ORIGIN.  
CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC  
CC DOMAIN. BELONGS TO THE SRC SUBFAMILY.  
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U00803; AAA18284.1; -  
DR EMBL; U2322; AAC50116.1; -  
DR HSP; P00523; 2PTK.

```
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_Kin.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00452; SH3DOMAIN.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYRKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR Transferase; Tyrosine-protein kinase; ATP-binding; SH2 domain;
KW SH3 domain; Phosphorylation.
FT DOMAIN 42 110 SH3.
FT DOMAIN 116 208 SH2.
FT DOMAIN 234 491 PROTEIN KINASE.
FT NP_BIND 240 248 ATP (BY SIMILARITY).
FT BINDING 262 262 ATP (BY SIMILARITY).
FT ACT_SITE 354 354 BY SIMILARITY.
FT MOD_RES 387 387 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT VARIANT 122 122 G -> R.
FT CONFLICT 115 115 P -> A (IN REF. 2).
FT SEQUENCE 505 AA; 58254 MW; 06EC050DBDCD930B CRC64;

Query Match 54.1%; Score 33; DB 1; Length 505;
Best Local Similarity 62.5%; Pred. No. 1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PQQFFGLM 11
||||:|
Db 459 PQQFVNIM 466

RESULT 50
SYK_ACICA
ID SYK_ACICA STANDARD; PRT; 509 AA.
AC Q43990;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE) (LYSRS).
GN LYSS.
OS Acinetobacter calcoaceticus.
OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC Acinetobacter.
OX NCBI_TaxID=471;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BD413 / ADP1.
RX MEDLINE=97228433; PubMed=9074511;
RA Gelsdoerfer W., Ratajczak A., Hillen W.;
RT "Nucleotide sequence of a putative periplasmic Mn superoxide dismutase
from Acinetobacter calcoaceticus ADP1.";
RL Gene 186:305-308(1997).
CC -1- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) -> AMP +
PYROPHOSPHATE + L-LYSYL-TRNA(LYS).
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
DR EMBL; Z46863; CAA86924.1; -.
DR HSP; P14825; ILYL.
DR InterPro; IPR002106; AA_trna_ligase_II.
DR InterPro; IPR002309; trna-synt_2.
DR InterPro; IPR002313; trna-synt_lys_2.
DR Pfam; PF00152; trna-synt_2; 1.
DR Pfam; PF01336; trna-anti_1.
DR PRINTS; PR00982; TRNASYNTHLYS.
DR PROSITE; PS00179; AA-trna_ligase_II_1; 1.
DR PROSITE; PS00339; AA-trna_ligase_II_2; 1.
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding.
SQ SEQUENCE 509 AA; 58079 MW; 95ED1AA43DC3D2F6 CRC64;

Query Match 54.1%; Score 33; DB 1; Length 509;
Best Local Similarity 60.0%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPQFFGL 10
||||:|
Db 153 RPLDPKFHGL 162

RESULT 51
COX1_APILI
ID COX1_APILI STANDARD; PRT; 521 AA.
AC P20374;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE I (EC 1.9.3.1).
GN COI.
OS Apis mellifera ligustica (Common honeybee).
OG Mitochondrion.
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata;
OC Apoidea; Apidae; Apis.
OX NCBI_TaxID=7469;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thorax;
RX MEDLINE=90136028; PubMed=2559293;
RA Crozier R.H., Crozier Y.C., Mackinlay A.G.;
RT "The CO-I and CO-II region of honeybee mitochondrial DNA: evidence
for variation in insect mitochondrial evolutionary rates.";
RL Mol. Biol. Evol. 6:399-411(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Thorax;
RX MEDLINE=93114603; PubMed=8417993;
RA Crozier R.H., Crozier Y.C.;
RT "The mitochondrial genome of the honeybee Apis mellifera: complete
sequence and genome organization.";
RL Genetics 133:97-117(1993).
CC -1- FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY
CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNITS 1-
3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. CO I IS THE
CATALYTIC SUBUNIT OF THE ENZYME. ELECTRONS ORIGINATING IN
CYTOCHROME C ARE TRANSFERRED VIA THE COPPER A CENTER OF SUBUNIT 2-
AND HEME A OF SUBUNIT 1 TO THE BIMETALLIC CENTER FORMED BY HEME A3
AND COPPER B.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) -> 2 H(2)O +
4 FERRICYTOCHROME C.
CC -1- PATHWAY: TERMINAL STEP IN THE RESPIRATORY CHAIN.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. MITOCHONDRIAL
INNER MEMBRANE. CONTAINS 12 POTENTIAL TRANSMEMBRANE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE HEME-COPPER RESPIRATORY OXIDASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
CC EMBL; M23409; AAL18476.1; -;  
CC EMBL; L06178; AAB96799.1; -;  
CC PIR; A32431; A32431.  
CC HSP; P00396; 10CC.  
CC InterPro; IPR000883; COX1.  
CC Pfam; PF00115; COX1; 1.  
CC PRINTS; PR01165; CYCOXIDASE1.  
CC PROSITE; PS00077; COX1; 1.  
CC Oxidoreductase; Heme; Copper; Mitochondrion; Transmembrane;  
KW Respiratory chain; Inner membrane.  
FT METAL 59  
FT METAL 59 IRON (HEME A) (PROBABLE).  
FT METAL 238 COPPER B (PROBABLE).  
FT METAL 242 COPPER B (PROBABLE).  
FT METAL 288 COPPER B (PROBABLE).  
FT METAL 289 COPPER B (PROBABLE).  
FT METAL 374 IRON (HEME A3) (PROBABLE).  
FT METAL 376 IRON (HEME A) (PROBABLE).  
SQ SEQUENCE 521 AA; 59293 MW; 2149417AC981CE64 CRC64;

Query Match 54.1%; Score 33; DB 1; Length 521;  
Best Local Similarity 75.0%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQOFFGLM 11  
II I III  
DB 425 PQHFLGLM 432

RESULT 52  
ID IMALARATH STANDARD; PRT; 596 AA.  
AC Q96321;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE IMPORTIN ALPHA-1 SUBUNIT (KARYOPHERIN ALPHA-1 SUBUNIT) (KAP ALPHA).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Ballas N., Citovsky V.;  
RT "ATKAPalpha gene from Arabidopsis encodes a protein that mediates nuclear import of Agrobacterium VirD2 protein.";  
RL (In) Plant Gene Register PGR97-129.  
CC -!- FUNCTION: BINDS SPECIFICALLY AND DIRECTLY TO SUBSTRATES CONTAINING EITHER A SIMPLE OR BIPARTITE NLS MOTIF. PROMOTES DOCKING OF IMPORT SUBSTRATES TO THE NUCLEAR ENVELOPE. SEEMS TO ACT AS A CYTOSOLIC RECEPTOR FOR BOTH SIMPLE AND BIPARTITE NLS MOTIFS (BY SIMILARITY).  
CC CELLULAR RECEPTOR FOR THE NUCLEAR IMPORT OF THE VIRD2 PROTEIN OF AGROBACTERIUM.  
CC -!- SUBUNIT: FORMS A COMPLEX WITH IMPORTIN BETA-1 SUBUNIT (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE IMPORTIN ALPHA FAMILY.  
CC -!- SIMILARITY: CONTAINS 8 ARM REPEATS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
CC EMBL; U69533; AAB72116.1; -;  
CC HSP; Q02248; 2BCT.  
CC InterPro; IPR000225; Armadillo.  
CC InterPro; IPR002652; IBB.  
CC Pfam; PF01749; IBB; 1.  
CC SMART; SM00185; ARM; 8.  
CC PROSITE; PS50176; ARM\_REPEAT; 3.  
KW Transport; Protein transport; Repeat.  
FT DOMAIN 12.  
FT REPEAT 109 151 IBB.  
FT REPEAT 152 196 ARM 1.  
FT REPEAT 197 234 ARM 2.  
FT REPEAT 235 279 ARM 3.  
FT REPEAT 280 319 ARM 4.  
FT REPEAT 320 362 ARM 5.  
FT REPEAT 363 403 ARM 6.  
FT REPEAT 403 445 ARM 7.  
FT REPEAT 446 596 ARM 8.  
FT DOMAIN 446 ASP/GLU-RICH (ACIDIC).  
SQ SEQUENCE 596 AA; 65606 MW; 2A2689E1C28FA3E7 CRC64;

Query Match 54.1%; Score 33; DB 1; Length 596;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKQOFF 8  
II I I I I I  
DB 231 RPKQPHF 238

RESULT 53  
TLD\_DROME

ID TLD\_DROME STANDARD; PRT; 1057 AA.  
AC P25723; Q9VC46;  
DT 01-MAY-1992 (Rel. 22, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DORSAL-VENTRAL PATTERNING TOLLID PROTEIN PRECURSOR (EC 3.4.24.-).  
GN TLD OR CG6868.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CANTON-S;  
RX MEDLINE=92034970; PubMed=1840509;  
RA Shmell M.J., Ferguson E.L., Childs S.R., O'Connor M.B.;  
RT "The Drosophila dorsal-ventral patterning gene tollid is related to human bone morphogenetic protein 1.";  
RL Cell 67:469-481(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95324373; PubMed=7600963;  
RA Finelli A.L., Bossie C.A., Xie T., Padgett R.W.;  
RT "Mutational analysis of the Drosophila tollid gene, a human BMP-1 homolog.";  
RL Development 120:861-870(1994).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X., Brannon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova K., Borchen M.R., Bouck J., Brokstein P., Brotier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahike C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulo G., Milshina N.V., Mobbart C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Rainert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RA "The genome sequence of *Drosophila melanogaster*."  
RL Science 287:2185-2195(2000).  
CC -I- FUNCTION: REQUIRED FOR NORMAL DORSAL DEVELOPMENT. TLD MAY INTERACT  
CC PHYSICALLY WITH DPP-C PROTEIN.  
CC -I- MISCELLANEOUS: MUTATIONS IN TLD GENE LEAD TO A PARTIAL  
CC TRANSFORMATION OF DORSAL ECTODERM INTO VENTRAL ECTODERM.  
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12A (ZINC  
CC METALLOPROTEASE); ALSO KNOWN AS THE ASTACIN SUBFAMILY.  
CC -I- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
CC -I- SIMILARITY: CONTAINS 5 CUB DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M76976; AAA28491.1; -;  
CC EMBL; U04239; AAC46482.1; -;  
CC EMBL; AE003749; AAF56329.1; -;  
CC PIR; A39288; A39288.  
CC HSP; P00742; IFAX.  
CC MEROPS; M12.010; -;  
CC FlyBase; FBgn0003719; tld.  
CC InterPro; IPR001506; Astacin.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000859; CUB.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR000130; Zn\_Mtpeptdse.  
DR Pfam; PF01400; Astacin; 1.  
DR Pfam; PF00431; CUB; 5.  
DR Pfam; PF00008; EGF; 2.  
DR PRINTS; PR00480; ASTACIN.  
DR SMART; SM00042; CUB; 5.  
DR SMART; SM00179; EGF\_Ca; 2.  
DR SMART; SM00235; ZnMc; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
DR PROSITE; PS01180; CUB; 5.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 2.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; 2.

DR PROSITE; PS01187; EGF\_CA; 2.  
KW Developmental protein; Hydrolase; Metalloprotease; Zinc; Glycoprotein;  
KW EGF-like domain; Calcium; Signal; Repeat; Zymogen.  
FT SIGNAL 1 27 POTENTIAL.  
FT PROPEP 28 126 POTENTIAL.  
FT CHAIN 127 1057 DORSAL-VENTRAL PATTERNING TOLLOID  
FT METALLOPROTEIN.  
FT DOMAIN 127 329 METALLOPROTEASE.  
FT DOMAIN 330 467 CUB 1.  
FT DOMAIN 468 580 CUB 2.  
FT DOMAIN 581 621 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 624 742 CUB 3.  
FT DOMAIN 743 783 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 787 899 CUB 4.  
FT DOMAIN 900 1016 CUB 5.  
FT METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 222 222 BY SIMILARITY.  
FT METAL 225 225 ZINC (CATALYTIC) (BY SIMILARITY).  
FT METAL 231 231 ZINC (CATALYTIC) (BY SIMILARITY).  
FT SITE 235 237 CELL ATTACHMENT SITE (POTENTIAL).  
FT SITE 315 317 CELL ATTACHMENT SITE (POTENTIAL).  
FT DISULFID 330 380 BY SIMILARITY.  
FT DISULFID 407 429 BY SIMILARITY.  
FT DISULFID 468 495 BY SIMILARITY.  
FT DISULFID 522 544 BY SIMILARITY.  
FT DISULFID 585 596 BY SIMILARITY.  
FT DISULFID 592 605 BY SIMILARITY.  
FT DISULFID 607 620 BY SIMILARITY.  
FT DISULFID 624 652 BY SIMILARITY.  
FT DISULFID 683 706 BY SIMILARITY.  
FT DISULFID 747 758 BY SIMILARITY.  
FT DISULFID 754 767 BY SIMILARITY.  
FT DISULFID 769 782 BY SIMILARITY.  
FT DISULFID 787 813 BY SIMILARITY.  
FT DISULFID 840 862 BY SIMILARITY.  
FT DISULFID 900 930 BY SIMILARITY.  
FT DISULFID 957 979 BY SIMILARITY.  
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 431 431 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 533 533 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 781 781 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 854 854 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT SEQUENCE 1057 AA; 120575 MW; 76F4B5ABE7996FBA CRC64;  
Query Match 54.1%; Score 33; DB 1; Length 1057;  
Best Local Similarity 36.4%; Pred. No. 2.1e+02;  
Matches 4; Conservative 5; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RPKPQQFFGLM 11  
: : : : :  
Db 42 KEQPEDFGIL 52  
RESULT 54  
ODPB\_BACST STANDARD; PRT; 324 AA.  
AC P21874;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-MAY-1991 (Rel. 18, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE PYRUVATE DEHYDROGENASE E1 COMPONENT, BETA SUBUNIT (EC 1.2.4.1).  
GN PDHB.  
OS Bacillus stearothermophilus.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Geobacillus.  
OX NCBI\_TaxID=1422;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=NCA 1503;



RX MEDLINE-90345939; PubMed-2200674;  
RA Hawkins C.F., Borges A., Perham R.N.;  
RT "Cloning and sequence analysis of the genes encoding the alpha and  
RT beta subunits of the E1 component of the pyruvate dehydrogenase  
RT multienzyme complex of Bacillus stearothermophilus.";  
RL Eur. J. Biochem. 191:337-346(1990).  
CC -!- FUNCTION: THE PYRUVATE DEHYDROGENASE COMPLEX CATALYZES THE OVERALL  
CC CONVERSION OF PYRUVATE TO ACETYL-COA & CO(2). IT CONTAINS MULTIPLE  
CC COPIES OF THREE ENZYMIC COMPONENTS: PYRUVATE DEHYDROGENASE (E1),  
CC DIHYDROLIPOAMIDE ACETYLTRANSFERASE (E2) & LIPOAMIDE DEHYDROGENASE  
CC (E3).  
CC -!- CATALYTIC ACTIVITY: PYRUVATE + LIPOAMIDE = S-ACETYL-DIHYDRO-  
CC LIPOAMIDE + CO(2).  
CC -!- COFACTOR: THIAMINE PYRROPHOSPHATE.  
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X53560; CAA37629.1; -  
DR PIR; S14230; S14230.  
DR InterPro; IPR000360; Transketolase.  
DR Pfam; PF00456; transketolase; 1.  
KW Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate.  
FT INIT\_MET 0  
SQ SEQUENCE 324 AA; 35328 MW; F7C6085E33371384 CRC64;

Query Match 53.3%; Score 32.5; DB 1; Length 324;  
Best Local Similarity 58.3%; Pred. No. 79;  
Matches 7; Conservative 2; Mismatches 2; Indels 1; Gaps 1;  
QY 1 RPQPO-QFFGLM 11  
DB 75 RPVPEIQFGFV 86  
11 1: 11111 1;  
11 1: 11111 1;

RESULT 55  
TKN\_PHYFU STANDARD; PRT; 11 AA.  
AC P08615;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE PHYSALAEAMIN.  
OS Physalaemus fuscumaculatus (Neotropical frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Leptodactylidae;  
OC Physalaemus.  
OX NCBI\_TaxID=8378;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE-66076612; PubMed-5857249;  
RA Ersparmer V., Anastasi A., Bertaccini G., Cei J.M.;  
RT "Structure and pharmacological actions of physalaemin, the main  
RT active polypeptide of the skin of Physalaemus fuscumaculatus.";  
RL Experientia 20:489-490(1964).  
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
DR PIR; S07201; S07201.  
DR InterPro; IPR003580; Protachykinin.  
DR InterPro; IPR002040; Tachykinin.  
DR Pfam; PF02202; Tachykinin; 1.  
DR SMART; SM00203; TK; 1.  
DR PROSITE; PS00267; TACHYKININ; 1.

KW Tachykinin; Neuropeptide; Amidation; Amphibian skin.  
FT MOD\_RES 1 1  
FT MOD\_RES 11 11  
SQ SEQUENCE 11 AA; 1283 MW; 3293693E59C33457 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 11;  
Best Local Similarity 62.5%; Pred. No. 3.3;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQFFGLM 11  
DB 4 PNKFYGLM 11  
1 1: 11111  
1 1: 11111

## RESULT 56

IBPL\_BOVIN STANDARD; PRT; 263 AA.  
ID IBPL\_BOVIN  
AC P24591;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE INSULIN-LIKE GROWTH FACTOR BINDING PROTEIN 1 PRECURSOR (IGFBP-1)  
DE (IBP-1) (IGF-BINDING PROTEIN 1).  
DE IGFBP1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-HOLSTEIN-FRIESIAN; TISSUE=Liver;  
RX MEDLINE-92119331; PubMed-1722724;  
RA Sneyers M., Kettmann R., Massart S., Renaville R., Burny A.,  
RA Portetelle D.;  
RT "Cloning and characterization of a cDNA encoding the bovine  
RT insulin-like growth factor binding protein 1 (biglycan-1).";  
RL DNA Seq. 1:407-408(1991).  
CC -!- FUNCTION: IGF-BINDING PROTEINS PROLONG THE HALF-LIFE OF THE IGFS  
CC AND HAVE BEEN SHOWN TO EITHER INHIBIT OR STIMULATE THE GROWTH  
CC PROMOTING EFFECTS OF THE IGFS ON CELL CULTURE. THEY ALTER THE  
CC INTERACTION OF IGFS WITH THEIR CELL SURFACE RECEPTORS.  
CC -!- SUBCELLULAR LOCATION: SECRETED.  
CC -!- MISCELLANEOUS: BINDS EQUALLY WELL IGF-I AND IGF-II.  
CC -!- SIMILARITY: CONTAINS 1 THYROGLOBULIN TYPE-I DOMAIN.  
CC -!- SIMILARITY: BELONGS TO THE INSULIN-LIKE GROWTH FACTOR BINDING  
CC PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X54979; CAA38723.1; -  
DR PIR; S23009; S23009.  
DR InterPro; IPR000867; IGFBP.  
DR InterPro; IPR000716; Thyroglobulin\_1.  
DR Pfam; PF00219; IGFBP; 1.  
DR Pfam; PF00086; thyroglobulin\_1; 1.  
DR SMART; SM00121; IB; 1.  
DR SMART; SM00211; TY; 1.  
DR PROSITE; PS00222; IGF\_BINDING; 1.  
DR PROSITE; PS00484; THYROGLOBULIN\_1; 1.  
KW Growth factor binding; Signal.  
FT SIGNAL 1 25  
FT CHAIN 26 263  
FT DOMAIN 206 255  
FT SITE 250 252

INSULIN-LIKE GROWTH FACTOR BINDING  
PROTEIN 1.  
THYROGLOBULIN TYPE I.  
CELL ATTACHMENT SITE.

FT DISULFID 73 86 BY SIMILARITY.  
 FT DISULFID 80 106 BY SIMILARITY.  
 FT DISULFID 180 210 BY SIMILARITY.  
 SQ SEQUENCE 263 AA; 28796 MW; 0403B642DDC45B6 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 263;  
 Best Local Similarity 66.7%; Pred. No. 80;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KPQOFFGL 10  
 |||||  
 Db 253 PKCQQYFNL 261

## RESULT 57

TRPA\_THETH STANDARD; PRT; 271 AA.  
 AC P16608;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE TRYPTOPHAN SYNTHASE ALPHA CHAIN (EC 4.2.1.20).  
 GN TRPA.  
 OS Thermus aquaticus (subsp. thermophilus).  
 OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.  
 OX NCBI\_TaxID=274;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HB27;  
 RX MEDLINE=90264352; PubMed=2189962;  
 RA Koyama Y., Furukawa K.;  
 RT "Cloning and sequence analysis of tryptophan synthetase genes of an  
 extreme thermophile, Thermus thermophilus HB27: plasmid transfer from  
 RT replica-plated Escherichia coli recombinant colonies to competent T.  
 RT thermophilus cells.";  
 RL J. Bacteriol. 172:3490-3495(1990).  
 CC -1- FUNCTION: THE ALPHA SUBUNIT IS RESPONSIBLE FOR THE ALDOL CLEAVAGE  
 CC OF INDOLGLYCEROL PHOSPHATE TO INDOL AND GLYCERALDEHYDE 3-  
 CC PHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: L-SERINE + 1-(INDOL-3-YL)GLYCEROL 3-PHOSPHATE  
 CC = L-TRYPTOPHAN + GLYCERALDEHYDE 3-PHOSPHATE + H(2)O.  
 CC -1- PATHWAY: LAST (FIFTH) STEP IN BIOSYNTHESIS OF TRYPTOPHAN.  
 CC -1- SUBUNIT: Tetramer of two alpha and two beta chains (by  
 CC similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TRPA FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; M32108; AAA27509.1; -  
 DR PIR: B35407; B35407.  
 DR HSSP: P00929; 1A5B.  
 DR InterPro: IPR002028; TRP\_synthase\_alpha.  
 DR Pfam: PF00290; trp\_synta; 1.  
 DR ProDom: PD001535; TRP\_synthase\_alpha; 1.  
 DR PROSITE: PS00167; TRP\_SYNTHASE\_ALPHA; 1.  
 KW Tryptophan biosynthesis; Lyase.  
 SQ SEQUENCE 271 AA; 28924 MW; C9E2A86080224DA2 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 271;  
 Best Local Similarity 71.4%; Pred. No. 82;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 PQOFFGL 10  
 |||||  
 Db 108 PERFFGL 114

## RESULT 58

YDHH\_HAEIN STANDARD; PRT; 382 AA.  
 AC P44861;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE HYPOTHETICAL PROTEIN HI0753.  
 GN HI0753.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OX Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RD / KW20 / ATCC 51907;  
 RX MEDLINE=95350630; PubMed=7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 RT influenzae Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- SIMILARITY: BELONGS TO THE UPF0075 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

EMBL; U32759; AAC22412.1; -  
 DR TIGR; HI0753; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 382 AA; 42026 MW; D14E353287BC11A6 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 382;  
 Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQOFFGLM 11  
 ||| : ||  
 Db 5 KPQYILGMM 13

## RESULT 59

GCH2\_CHLPN STANDARD; PRT; 418 AA.  
 AC Q92734; Q9JQ68;  
 DT 20-AUG-2001 (Rel. 40, Created)  
 DT 20-AUG-2001 (Rel. 40, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE RIBOFLAVIN BIOSYNTHESIS PROTEIN RIBA [INCLUDES: GTP CYCLOHYDROLASE II  
 DE (EC 3.5.4.25); 3,4-DIHYDROXY-2-BUTANONE 4-PHOSPHATE SYNTHASE (DHBP  
 DE SYNTHASE)].  
 GN RIBAB OR CPN0872 OR CP0997.  
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.   
 OX NCBI\_TaxID=83556;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CWL029;

RX MEDLINE=99206606; PubMed=10192388;  
RA Olman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,  
RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;  
RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis";  
RL Nat. Genet. 21:385-389(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=AR39;  
RX MEDLINE=20150255; PubMed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,  
RA Linhorst M., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,  
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,  
RA Eisen J., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia  
RT pneumoniae AR39";  
RL Nucleic Acids Res. 28:1397-1406(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=J138;  
RX MEDLINE=20330349; PubMed=10871362;  
RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,  
RA Shiba T., Ishii K., Hattori M., Kuhara S., Nakazawa T.;  
RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138  
RT from Japan and CWL029 from USA";  
RL Nucleic Acids Res. 28:2311-2314(2000).  
CC -1- CATALYTIC ACTIVITY: GTP + 3 H(2)O = FORMATE + 2,5-DIAMINO-6-  
CC HYDROXY-4-(5-PHOSPHORIBOSYLAMINO)PYRIMIDINE + PYROPHOSPHATE.  
CC -1- PATHWAY: RIBOFLAVIN BIOSYNTHESIS.  
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE DHP  
CC SYNTHASE FAMILY.  
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE GTP  
CC CYCLOHYDROLASE II FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AE001667; AAD19010.1; -;  
DR EMBL; AE002257; AAF38775.1; -;  
DR EMBL; AF002548; BAA99080.1; -;  
DR TIGR; CP0997; -;  
DR InterPro; IPR000422; DHP\_synthase.  
DR InterPro; IPR000926; GTP\_cyclohydro2.  
DR Pfam; PF00926; DHP\_synthase; 1.  
DR Pfam; PF00925; GTP\_cyclohydro2; 1.  
DR ProDom; PD003034; DHP\_synthase; 1.  
KW Multifunctional enzyme; Riboflavin biosynthesis; Hydrolase;  
KW Complete proteome.  
FT DOMAIN 1 211 DHP SYNTHASE.  
FT DOMAIN 212 418 GTP CYCLOHYDROLASE II.  
SQ SEQUENCE 418 AA; 45845 MW; 7A5A214BB0EC0E32 CRC64;  
  
Query Match 52.5%; Score 32; DB 1; Length 418;  
Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 4 PQOFFGL 10  
DB 366 PQKIFGL 372  
  
RESULT 60  
RHO\_BORBU STANDARD; PRT; 419 AA.  
AC P33561; O51248;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)

20-AUG-2001 (Rel. 40, Last annotation update)  
DE TRANSCRIPTION TERMINATION FACTOR RHO.  
GN RHO OR BB0230.  
OS Borrelia burgdorferi (Lyme disease spirochete).  
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.  
OX NCBI\_TaxID=139;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93197131; PubMed=8451174;  
RA Tilly K., Campbell J.;  
RT "A Borrelia burgdorferi homolog of the Escherichia coli rho gene";  
RL Nucleic Acids Res. 21:1040-1040(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 35210 / B31;  
RX MEDLINE=98065943; PubMed=9403685;  
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,  
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,  
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,  
RA Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,  
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,  
RA Utterback T., Wathley L., McDonald L., Artiach P., Bowman C.,  
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,  
RA Smith H.O., Venter J.C.;  
RT "Genomic sequence of a Lyme disease spirochete, Borrelia  
RT burgdorferi";  
RL Nature 390:580-586(1997).  
RN [3]  
RP SEQUENCE OF 46-235 FROM N.A.  
RC STRAIN=212;  
RX MEDLINE=95111614; PubMed=7812434;  
RA Ojalimi C., Davidson B.E., Saint-Girons I., Old I.G.;  
RT "Conservation of gene arrangement and an unusual organization of rRNA  
RT genes in the linear chromosomes of the Lyme disease spirochaetes  
RT Borrelia burgdorferi, B. garinii and B. afzelii";  
RL Microbiology 140:2931-2940(1994).  
CC -1- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM  
CC THAT INVOLVES RHO BINDING TO THE NASCENT RNA. ACTIVATION OF RHO'S  
CC RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE MRNA FROM THE  
CC DNA TEMPLATE (BY SIMILARITY).  
CC -1- SUBUNIT: HOMOHXAMER (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L07656; AAA71920.1; ALT\_INIT.  
DR EMBL; U35673; ABA41463.1; ALT\_INIT.  
DR EMBL; AE001133; AAC66619.1; ALT\_INIT.  
DR EMBL; L46347; AAA73991.1; -;  
DR HSSP; P03002; IAG3.  
DR TIGR; BB0230; -;  
DR InterPro; IPR003593; AAA.  
DR InterPro; IPR000194; ATPase\_alpha\_beta.  
DR InterPro; IPR002059; Cold\_shock.  
DR Pfam; PF00006; ATP-synt\_ab; 1.  
DR SMART; SM00382; AAA; 1.  
DR SMART; SM00357; CSP; 1.  
KW Transcription termination; Helicase; ATP-binding; RNA-binding;  
KW Complete proteome.  
FT DOMAIN 21 26 RNA-BINDING (RNP2) (BY SIMILARITY).  
FT DOMAIN 63 66 RNA-BINDING (RNP1) (BY SIMILARITY).  
FT NP\_BIND 177 184 ATP (POTENTIAL).  
FT CONFLICT 58 58 D -> H (IN REF. 3).  
FT CONFLICT 210 210 E -> D (IN REF. 3).  
SQ SEQUENCE 419 AA; 46950 MW; 38D0C354E6C00ABF CRC64;

Query Match 52.5%; Score 32; DB 1; Length 419;  
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 KPOOFFG 9  
 ||::|||  
 Db 295 KPKRFFG 301

RESULT 61  
 RHO\_PSEFL STANDARD; PRT; 419 AA.  
 AC P52155;  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DE TRANSCRIPTION TERMINATION FACTOR RHO.  
 GN RHO.  
 OS Pseudomonas fluorescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=294;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 17400;  
 RX MEDLINE=94327472; PubMed=8051015;  
 RA Opperman T., Richardson J.P.;  
 RT "Phylogenetic analysis of sequences from diverse bacteria with  
 RT homology to the Escherichia coli rho gene.";  
 RL J. Bacteriol. 176:5033-5043(1994).  
 CC -!- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM  
 CC THAT INVOLVES RHO BINDING TO THE NASCENT RNA, ACTIVATION OF RHO'S  
 CC RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE MRNA FROM THE  
 CC DNA TEMPLATE (BY SIMILARITY).  
 CC -!- SUBUNIT: HOMOHXAMER (BY SIMILARITY).  
 CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L27278; AAA59209.1; -  
 CC HSP; P03002; IAE3.  
 CC InterPro; IPR003593; AAA.  
 CC InterPro; IPR000194; ATPase\_alpha\_beta.  
 CC Pfam; PF00006; ATP-synt\_ab; 1.  
 CC SMART; SM00382; AAA; 1.  
 CC SMART; SM00357; CSP; 1.  
 CC Transcription termination; Helicase; ATP-binding; RNA-binding.  
 KW DOMAIN 21 26 RNA-BINDING (RNP2) (BY SIMILARITY).  
 FT DOMAIN 61 64 RNA-BINDING (RNP1) (BY SIMILARITY).  
 FT NP\_BIND 178 185 ATP (POTENTIAL).  
 FT SEQUENCE 419 AA; 46954 MW; ED30BDE6ACE03253 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 419;  
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 KPOOFFG 9  
 ||::|||  
 Db 296 KPKRFFG 302

RESULT 62  
 RHO\_THEMA STANDARD; PRT; 427 AA.  
 ID RHO\_THEMA  
 AC P38527;

DT 01-OCT-1994 (Rel. 30, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE TRANSCRIPTION TERMINATION FACTOR RHO.  
 GN RHO OR TMI470.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogales; Thermotoga.  
 OX NCBI\_TaxID=2336;  
 RN [1]  
 RP SEQUENCE OF 16-427 FROM N.A.  
 RX MEDLINE=94327472; PubMed=8051015;  
 RA Opperman T., Richardson J.P.;  
 RT "Phylogenetic analysis of sequences from diverse bacteria with  
 RT homology to the Escherichia coli rho gene.";  
 RL J. Bacteriol. 176:5033-5043(1994).  
 CC -!- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM  
 CC THAT INVOLVES RHO BINDING TO THE NASCENT RNA, ACTIVATION OF RHO'S  
 CC RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE MRNA FROM THE  
 CC DNA TEMPLATE (BY SIMILARITY).  
 CC -!- SUBUNIT: HOMOHXAMER (BY SIMILARITY).  
 CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L27279; AAA59210.1; -  
 CC EMBL; AE001798; AAD36538.1; -  
 CC HSP; P03002; IAE2.  
 CC TIGR; TMI470; -  
 CC InterPro; IPR003593; AAA.  
 CC InterPro; IPR000194; ATPase\_alpha\_beta.  
 CC InterPro; IPR002059; Cold\_shock.  
 CC Pfam; PF00006; ATP-synt\_ab; 1.  
 CC SMART; SM00382; AAA; 1.  
 CC SMART; SM00357; CSP; 1.  
 CC Transcription termination; Helicase; ATP-binding; RNA-binding;  
 KW Complete proteome.  
 FT DOMAIN 28 33 RNA-BINDING (RNP2) (BY SIMILARITY).  
 FT DOMAIN 68 71 RNA-BINDING (RNP1) (BY SIMILARITY).  
 FT NP\_BIND 182 189 ATP (POTENTIAL).  
 FT SEQUENCE 427 AA; 48301 MW; 37748653910AFC95 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 427;  
 Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 KPOOFFG 9  
 ||::|||  
 Db 300 KPKRFFG 306

RESULT 63  
 MM12\_MOUSE STANDARD; PRT; 462 AA.  
 ID MM12\_MOUSE

P34960;  
01-FEB-1994 (Rel. 28, Created)  
01-FEB-1994 (Rel. 28, Last sequence update)  
20-AUG-2001 (Rel. 40, Last annotation update)  
MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX  
DE METALLOPROTEINASE-12) (MMP-12).  
GN MMP12 OR MMEL OR MME.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 99-125.  
RC TISSUE=Macrophage;  
RX MEDLINE=92165826; PubMed=1537850;  
RA Shapiro S.D., Griffin G.L., Gilbert D.J., Jenkins N.A.,  
RA Copeland N.G., Welgus H.G., Senior R.M., Ley T.J.;  
RT "Molecular cloning, chromosomal localization, and bacterial  
expression of a murine macrophage metalloelastase";  
RL J. Biol. Chem. 267:4664-4671(1992).  
CC -!- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS  
CC SIGNIFICANT ELASTOLYTIC ACTIVITY.  
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.  
CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-  
CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.  
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.  
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M82831; AAA39526.1; -;  
CC F1R; A42401; A42401.  
CC HSSP; P03956; ICGL.  
CC MEROPS; M10.009; -;  
CC MGD; MGI:97005; Mmp12.  
CC InterPro; IPR000585; Hemopexin.  
CC InterPro; IPR001818; Matrixin.  
CC InterPro; IPR000130; Zn\_MTPeptase.  
CC Pfam; PF00045; hemopexin; 4.  
CC PRINTS; PR00138; MATRIXIN.  
CC SMART; SM00120; HX; 4.  
CC SMART; SM00235; ZnMc; 1.  
CC PROSITE; PS00024; HEMOPEXIN; 1.  
CC PROSITE; PS00142; ZINC\_PROTEASE; 1.  
CC PROSITE; PS00546; CYSTEINE\_SWITCH; FALSE\_NEG.  
CC Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;  
CC Extracellular matrix; Signal.  
CC SIGNAL 1 17 PROBABLE.  
CC PROPEP 18 98 ACTIVATION PEPTIDE.  
CC CHAIN 99 462 MACROPHAGE METALLOELASTASE.  
CC DOMAIN 272 462 HEMOPEXIN-LIKE.  
CC SITE 85 85 CYSTEINE SWITCH (BY SIMILARITY).  
CC METAL 211 211 ZINC (CATALYTIC) (BY SIMILARITY).  
CC ACT\_SITE 212 212 BY SIMILARITY.  
CC METAL 215 215 ZINC (CATALYTIC) (BY SIMILARITY).  
CC METAL 221 221 ZINC (CATALYTIC) (BY SIMILARITY).  
CC CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT CARBOHYD 310 310 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC FT DISULFID 275 462 BY SIMILARITY.  
CC SEQUENCE 462 AA; 53841 MW; BB9625906FIDBDF CRC64;  
Query Match 52.5%; Score 32; DB 1; Length 462;

Best Local Similarity 100.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 5 QQFFGL 10  
DB 61 QQFFGL 66  
RESULT 64  
MM12\_RAT  
ID MM12\_RAT STANDARD; PRT; 465 AA.  
AC Q63341;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MACROPHAGE METALLOELASTASE PRECURSOR (EC 3.4.24.65) (MME) (MATRIX  
DE METALLOPROTEINASE-12) (MMP-12).  
GN MMP12 OR MMEL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RA Cossins J., Clements J., Catlin G.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: MAY BE INVOLVED IN TISSUE INJURY AND REMODELING. HAS  
CC SIGNIFICANT ELASTOLYTIC ACTIVITY (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF SOLUBLE AND INSOLUBLE ELASTIN.  
CC SPECIFIC CLEAVAGES ARE ALSO PRODUCED AT 14-ALA-|-LEU-15 AND 16-  
CC TYR-|-LEU-17 IN THE B CHAIN OF INSULIN.  
CC -!- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY (BY SIMILARITY).  
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X98517; CAA67142.1; -;  
CC HSSP; P03956; ICGL.  
CC MEROPS; M10.009; -;  
CC InterPro; IPR000585; Hemopexin.  
CC InterPro; IPR001818; Matrixin.  
CC InterPro; IPR000130; Zn\_MTPeptase.  
CC Pfam; PF00045; hemopexin; 4.  
CC PRINTS; PR00138; MATRIXIN.  
CC SMART; SM00120; HX; 4.  
CC SMART; SM00235; ZnMc; 1.  
CC PROSITE; PS00024; HEMOPEXIN; 1.  
CC PROSITE; PS00142; ZINC\_PROTEASE; 1.  
CC PROSITE; PS00546; CYSTEINE\_SWITCH; FALSE\_NEG.  
CC Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;  
CC Extracellular matrix; Signal.  
CC SIGNAL 1 21 PROBABLE.  
CC PROPEP 22 101 ACTIVATION PEPTIDE (BY SIMILARITY).  
CC CHAIN 102 465 MACROPHAGE METALLOELASTASE.  
CC DOMAIN 275 465 HEMOPEXIN-LIKE.  
CC SITE 88 88 CYSTEINE SWITCH (BY SIMILARITY).  
CC METAL 214 214 ZINC (CATALYTIC) (BY SIMILARITY).  
CC ACT\_SITE 215 215 BY SIMILARITY.  
CC METAL 218 218 ZINC (CATALYTIC) (BY SIMILARITY).  
CC METAL 224 224 ZINC (CATALYTIC) (BY SIMILARITY).  
CC CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).  
CC DISULFID 278 465 BY SIMILARITY.  
CC SEQUENCE 465 AA; 53738 MW; E779B6Q14EC6FF68 CRC64;

```
Query Match      52.5%; Score 32; DB 1; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QOFFGL 10
Db 64 QOFFGL 69

RESULT 65
MM01_PIG STANDARD; PRT; 469 AA.
AC P21692;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE INTERSTITIAL COLLAGENASE PRECURSOR (BC 3.4.24.7) (MATRIX
DE METALLOPROTEINASE-1) (MMP-1).
GN MMP1.
OS Sus. scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91333421; PubMed=1651440;
RA Richards C.D., Rafferty J.A., Reynolds J.J., Saklatvala J.;
RT "Porcine collagenase from synovial fibroblasts: cDNA sequence and
RT modulation of expression of RNA in vitro by various cytokines.";
RL Matrix 11:161-167(1991).
RN [2]
RP SEQUENCE OF 25-469 FROM N.A.
RC TISSUE=Synovial cell;
RX MEDLINE=91067477; PubMed=2174547;
RA Clarke N.J., O'Hare M.C., Cawston T.E., Harper G.P.;
RT "Nucleotide sequence of a cDNA for porcine type I collagenase,
RT obtained by PCR.";
RL Nucleic Acids Res. 18:6703-6703(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 100-469.
RX MEDLINE=96173003; PubMed=8590015;
RA Li J., Brick P., O'Hare M.C., Skarzynski T., Lloyd L.F., Curry V.A.,
RA Clark I.M., Bigg H.F., Hazleman B.L., Cawston T.E., Blow D.M.;
RT "Structure of full-length porcine synovial collagenase reveals a C-
RT terminal domain containing a calcium-linked, four-bladed
RT beta-propeller.";
RL Structure 3:541-549(1995).
RN [4]
RP SEQUENCE OF 100-104 AND 248-282, AND AUTOLYTIC CLEAVAGE SITE.
RX MEDLINE=95142615; PubMed=7840605;
RA Clark I.M., Mitchell R.E., Powell L.K., Bigg H.F., Cawston T.E.,
RA O'Hare M.C.;
RT "Recombinant porcine collagenase: purification and autolysis.";
RL Arch. Biochem. Biophys. 316:123-127(1995).
CC -!- FUNCTION: CLEAVES COLLAGENS OF TYPES I, II, AND III AT ONE SITE IN
CC THE HELICAL DOMAIN. ALSO CLEAVES COLLAGENS OF TYPES VII AND X.
CC -!- COPACITOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -!- ENZYME REGULATION: CAN BE ACTIVATED WITHOUT REMOVAL OF THE
CC ACTIVATION PEPTIDE.
CC -!- PTM: UNDERGOES AUTOLYTIC CLEAVAGE TO PRODUCE A N-TERMINAL
CC FRAGMENT HAVING REDUCED COLLAGENOLYTIC ACTIVITY.
CC -!- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEINASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/)
-----
or send an email to license@isb-sib.ch.
-----
EMBL; X54724; CAA38526.1; -.
PIR; SI5986; KCFGI.
PDB; IFBL; 29-JAN-96.
MEROPS; M10.001; -.
DR InterPro; IPR000585; Hemoexin.
DR InterPro; IPR001818; Matrixin.
DR InterPro; IPR000130; Zn_Mtpeptidse.
Pfam; PF00045; hemoexin; 4.
Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
KW Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
KW Collagen degradation; Extracellular matrix; Autocatalytic cleavage;
KW Signal; 3D-structure.
FT SIGNAL 1 19 ACTIVATION PEPTIDE.
FT PROPEP 20 99 INTERSTITIAL COLLAGENASE.
FT CHAIN 100 469 18 KDA INTERSTITIAL COLLAGENASE (WEAK
FT CHAIN 100 258 COLLAGENASE ACTIVITY).
FT DOMAIN 275 469 HEMOPEXIN-LIKE.
FT SITE 92 92 CYSTEINE SWITCH (POTENTIAL).
FT SITE 258 259 CLEAVAGE (AUTOLYTIC).
FT METAL 218 218 ZINC (CATALYTIC).
FT ACT_SITE 219 219 ZINC (CATALYTIC).
FT METAL 222 222 ZINC (CATALYTIC).
FT METAL 228 228 PROBABLE.
FT DISULFID 278 466 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 120 120
SQ SEQUENCE 469 AA; 53666 MW; 7952D7B2753F682 CRC64;

Query Match      52.5%; Score 32; DB 1; Length 469;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QOFFGL 10
Db 68 QOFFGL 73

RESULT 66
SYK_BACST STANDARD; PRT; 494 AA.
AC Q9RHV9;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE) (LYSRS).
GN LYSS.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Geobacillus.
OX NCBI_TaxID=1422;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NCA 1503;
RX MEDLINE=20199468; PubMed=10737207;
RA Takita T., Shimizu N., Sukata T., Hashimoto S., Akita E., Yokota T.,
RA Esaki N., Soda K., Inouye K., Tomomura B.;
RT "Lysyl-tRNA synthetase of Bacillus stearothermophilus molecular
RT cloning and expression of the gene.";
RL Biosci. Biotechnol. Biochem. 64:432-437(2000).
CC -!- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) = AMP +
CC PYROPHOSPHATE + L-LYSYL-TRNA(LYS).
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.
```

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AB012100; BAA88691.1; -  
DR InterPro; IPR002106; AA\_trna\_ligase\_II.  
DR InterPro; IPR002309; trna-synt\_2.  
DR InterPro; IPR002312; trna-synt\_asp.  
DR InterPro; IPR002313; trna-synt\_lys\_2.  
DR Pfam; PF00152; trna-synt\_2; 1.  
DR Pfam; PF01336; trna-anti\_1.  
DR PRINTS; PRO0982; TRNASYNTHLVS.  
DR PROSITE; PS01042; TRNASYNTHASP.  
DR PROSITE; PS00179; AA\_TRNA\_LIGASE\_II.1; 1.  
DR PROSITE; PS00339; AA\_TRNA\_LIGASE\_II.2; 1.  
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
SQ SEQUENCE 494 AA; 57405 MW; 109D1A4FDD7F714C CRC64;  
-----  
Query Match 52.5%; Score 32; DB 1; Length 494;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RPKQQQFFGL 10  
|||::||  
Db 142 RPLPERYHGL 151  
-----  
RESULT 67  
RHO\_TREPA  
ID RHO\_TREPA STANDARD; PRT; 519 AA.  
AC 083281;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE TRANSCRIPTION TERMINATION FACTOR RHO.  
GN RHO OR TP0254.  
OS Treponema pallidum.  
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.  
OX NCBI\_TaxID=160;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NICHOLS;  
RX MEDLINE=98332770; PubMed=9665876;  
RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G., Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A., Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J., Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T., McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S., Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O., Venter J.C.;  
RT "Complete genome sequence of Treponema pallidum, the syphilis spirochete.";  
RL Science 281:375-388(1998).  
CC -!- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM THAT INVOLVES RHO BINDING TO THE NASCENT RNA, ACTIVATION OF RHO'S RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE MRNA FROM THE DNA TEMPLATE (BY SIMILARITY).  
CC -!- SUBUNIT: HOMOHETEROMER (BY SIMILARITY).  
CC -!- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

CC EMBL; AE001207; AAC65243.1; -  
DR TIGR; TP0254; -  
DR InterPro; IPR003593; AAA.  
DR InterPro; IPR000194; ATPase\_alpha\_beta.  
DR InterPro; IPR002059; Cold\_shock.  
DR Pfam; PF00006; ATP-synt\_ab; 1.  
DR SMART; SM00382; AAA; 1.  
DR SMART; SM00357; CSP; 1.  
KW Transcription termination; Helicase; ATP-binding; RNA-binding;  
KW Complete proteome.  
FT DOMAIN 118 125 RNA-BINDING (RNP2) (BY SIMILARITY).  
FT DOMAIN 160 163 RNA-BINDING (RNP1) (BY SIMILARITY).  
FT NP\_BIND 274 281 ATP (POTENTIAL).  
SQ SEQUENCE 519 AA; 58265 MW; 321A637776025A10 CRC64;  
-----  
Query Match 52.5%; Score 32; DB 1; Length 519;  
Best Local Similarity 71.4%; Pred. No. 1.6e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 3 KPOQFFG 9  
|||::|||  
Db 392 KPKRFFG 398  
-----  
RESULT 68  
RAI2\_MOUSE  
ID RAI2\_MOUSE STANDARD; PRT; 529 AA.  
AC Q9QVY8;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE RETINOIC ACID-INDUCED PROTEIN 2 (3F8).  
GN RAI2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94148137; PubMed=8314004;  
RA Jonk L.J.C., de Jonge M.E., Vervaaert J.M., Wissink S., Kruijer W.;  
RT "Isolation and developmental expression of retinoic-acid-induced genes.";  
RT Dev. Biol. 161:604-614(1994).  
CC -!- INDUCTION: BY RETINOIC ACID.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X76652; -; NOT\_ANNOTATED\_CDS.  
DR MGD; MGI:1344378; Rai2.  
FT DOMAIN 200 253 PRO-RICH.  
SQ SEQUENCE 529 AA; 57178 MW; 10AA48B170FCDBD0 CRC64;  
-----  
Query Match 52.5%; Score 32; DB 1; Length 529;  
Best Local Similarity 66.7%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 PKPQQFFGL 10  
|||::|||  
Db 265 PKPPSSFFGL 273  
-----  
RESULT 69  
RAI2\_HUMAN

```

ID RA12_HUMAN STANDARD; PRT; 530 AA.
AC Q9Y5P3;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE RETINOIC ACID-INDUCED PROTEIN 2.
GN RA12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99168896; PubMed=10049581;
RA Walpole S.M., Hiriyana K.T., Nicolau A., Bingham E.L., Durham J.,
RA Vaudin M., Ross M.T., Yates J.R., Steving P.A., Trump D.;
RT "Identification and characterization of the human homologue (RA12) of
RT a mouse retinoic acid-induced gene in Xp22.2."
RL Genomics 55:275-283(1999).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF136587; AAD33688.1; -.
DR MIM; 300217; -.
DR DOMAIN 200 253 PRO-RICH.
FT SEQUENCE 530 AA; 57148 MW; 9879EE869DC6188F CRC64;
SQ
Query Match 52.5%; Score 32; DB 1; Length 530;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Oy 2 PKPQOFFGL 10
|||
Db 265 PKPPSFFGL 273
-----
RESULT 70
ID C4C3_DROME STANDARD; PRT; 535 AA.
AC Q9VA27; Q24121;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CYTOCHROME P450 4C3 (EC 1.14.-.-) (CYP1VC3).
GN CYP4C3 OR CG1438.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RX STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
```

```

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Llang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
[2]
RP SEQUENCE OF 347-472 FROM N.A.
RC STRAIN-HAAG-79;
RX MEDLINE=96262181; PubMed=8676871;
RA Dunkov B.C., Rodriguez-Arnaiz R., Pittendrigh B.,
RA French-Constant R.H., Feyerisen R.;
RT "Cytochrome P450 gene clusters in Drosophila melanogaster."
RL Mol. Gen. Genet. 251:290-297(1996).
CC -!- FUNCTION: MAY BE INVOLVED IN THE METABOLISM OF INSECT HORMONES AND
CC IN THE BREAKDOWN OF SYNTHETIC INSECTICIDES (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) -> ROH +
CC OXIDIZED FLAVOPROTEIN + H(2)O.
CC -!- SUBCELLULAR LOCATION: MEMBRANE-BOUND. ENDOPLASMIC RETICULUM
CC (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE003775; AAP57098.1; -.
DR EMBL; U34323; AAA80657.1; -.
DR FlyBase; FBgn0015032; Cyp4c3.
DR InterPro; IPR001128; Cyt_P450.
DR Pfam; PF00067; p450; 1.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME_P450; 1.
KW Oxidoreductase; Monooxygenase; Membrane; Heme; Microsome;
KW Endoplasmic reticulum.
FT BINDING 481 481 HEME (BY SIMILARITY).
SQ SEQUENCE 535 AA; 60757 MW; 0C78200AC2D35979 CRC64;
Query Match 52.5%; Score 32; DB 1; Length 535;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Oy 2 PKPQQF 7
|||
Db 448 PKPEQF 453
-----
RESULT 71
```





SQ SEQUENCE 766 AA; 82363 MW; 1F3EFFD6DC84B100 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 766;

Best Local Similarity 75.0%; Pred. No. 2.3e+02; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKQOQFF 8  
|||||

Db 267 RPKQOASF 274

RESULT 73

DSC3\_BOVIN STANDARD; PRT; 896 AA.

AC Q28060; Q28176;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE DESMOCOLLIN 3A/3B PRECURSOR.

GN DSC3.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=95403557; PubMed=7673337;

RA Yue K.K.M., Holton J.L., Clarke J.P., Hyam J.L.M., Hashimoto T.,

RA Chidgey M.A.J., Garrod D.R.;

RT "Characterisation of a desmocollin isoform (bovine DSC3) exclusively

RT expressed in lower layers of stratified epithelia.";

RL J. Cell Sci. 108:2163-2173(1995).

RN [2]

RN SEQUENCE OF 686-814 FROM N.A.

RP TISSUE=Epidermis;

RX MEDLINE=94308280; PubMed=8034749;

RA Legan P.K., Yue K.K.M., Chidgey M.A.J., Holton J.L., Wilkinson R.W.,

RA Garrod D.R.;

RT "The bovine desmocollin family: a new gene and expression patterns

RT reflecting epithelial cell proliferation and differentiation.";

RL J. Cell Biol. 126:507-518(1994).

CC -1- FUNCTION: COMPONENT OF INTERCELLULAR DESMOSOME JUNCTIONS. INVOLVED

CC IN THE INTERACTION OF PLAQUE PROTEINS AND INTERMEDIATE FILAMENTS

CC MEDIATING CELL-CELL ADHESION. MAY CONTRIBUTE TO EPIDERMAL CELL

CC POSITIONING (STRATIFICATION) BY MEDIATING DIFFERENTIAL

CC ADHESIVENESS BETWEEN CELLS THAT EXPRESS DIFFERENT ISOFORMS.

CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 3A (SHOWN HERE) AND 3B; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: STRATIFIED EPITHELIA ONLY (EPIDERMIS, TONGUE,

CC ESOPHAGUS AND RUMEN).

CC -1- DOMAIN: CALCIUM MAY BE BOUND BY THE CADHERIN-LIKE REPEATS

CC (POTENTIAL).

CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

CC EMBL; L33774; AAC41625.1; -

CC EMBL; L33774; AAC41626.1; -

CC EMBL; X75783; CAA53427.1; -

CC HSP; P09803; 1SDU.

CC InterPro: IPR002126; Cadherin.

CC Pfam: PF00028; cadherin; 5.

CC PRINTS; PR00205; CADHERIN.

CC SMART; SM00112; CA; 5.

DR PROSITE; PS00232; CADHERIN\_1; 3.  
DR PROSITE; PS0268; CADHERIN\_2; 5.  
KW Cell adhesion; Glycoprotein; Transmembrane; Repeat; Signal;  
KW Alternative splicing; Cytoskeleton; Calcium-binding.  
FT SIGNAL 1 26  
FT PROPEP 27 134  
FT CHAIN 135 896  
FT DOMAIN 135 690  
FT TRANSMEM 691 711  
FT DOMAIN 712 896  
FT DOMAIN 135 242  
FT DOMAIN 243 354  
FT DOMAIN 355 471  
FT DOMAIN 472 579  
FT DOMAIN 580 690  
FT CARBOHYD 165 165  
FT CARBOHYD 391 391  
FT CARBOHYD 546 546  
FT CARBOHYD 629 629  
FT VARSPLIC 832 839  
FT VARSPLIC 840 896  
FT CONFLICT 686 687  
SQ SEQUENCE 896 AA; 99687 MW; 8CC0C30A63FB0BD4 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 896;

Best Local Similarity 100.0%; Pred. No. 2.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 QFFGLM 11

Db 330 QFFGLM 335

RESULT 74

ID STA2\_MOUSE

AC Q9WVL2; Q64189; Q64250; Q64188;

DT 20-AUG-2001 (Rel. 40, Created)

DT 20-AUG-2001 (Rel. 40, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)

DE SIGNAL TRANSDUCER AND ACTIVATOR OF TRANSCRIPTION 2.

GN STAT2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RN SEQUENCE FROM N.A.

RP STRAIN=CD-1;

RA Paulson M.S., Mui A., Levy D.E.;

RT "Molecular cloning and characterization of murine Stat2.";

RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RN SEQUENCE OF 595-658 FROM N.A. (ISOFORMS A AND B/C).

RX MEDLINE=96176320; PubMed=8601453;

RA Sugiyama T., Nishio Y., Kishimoto T., Akira S.;

RT "Identification of alternative splicing form of Stat2.";

RL FEBS Lett. 381:191-194(1996).

CC -1- FUNCTION: TRANSCRIPTION FACTOR THAT BINDS TO THE IFN-STIMULATED

CC RESPONSE ELEMENT (ISRE) AND TO THE GAS ELEMENT. THIS MULTIPROTEIN

CC TRANSCRIPTION FACTOR IS TERMED ISGF3.

CC -1- SUBUNIT: IN RESPONSE TO IFN ALPHA/BETA, THREE SUBUNITS (STAT1-

CC ALPHA, STAT1-BETA, STAT2) OF ISGF3, BECOME PHOSPHORYLATED ON

CC TYROSINE, MIGRATE INTO THE NUCLEUS, AND ASSEMBLE INTO A COMPLEX

CC TOGETHER WITH ISGF3 GAMMA (P48), A DNA-BINDING PROTEIN THAT

CC SPECIFICALLY BINDS TO THE IFN-STIMULATED RESPONSE ELEMENT (BY

CC SIMILARITY).

CC -1- SUBCELLULAR LOCATION: NUCLEAR; TRANSLOCATED INTO THE NUCLEUS IN

CC RESPONSE TO PHOSPHORYLATION (BY SIMILARITY).

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B/C; ARE

CC PRODUCED BY ALTERNATIVE SPLICING.

CC -1- TISSUE SPECIFICITY: FOUND IN THE BRAIN, LUNG, HEART, SPLEEN,

CC LIVER, KIDNEY, MUSCLE, AND THE TESTIS.  
 CC -!- PTM: TYROSINE PHOSPHORYLATED IN RESPONSE TO IFN-ALPHA (BY  
 CC SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE STAT FAMILY OF TRANSCRIPTION FACTORS.  
 CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF088862; AAD38329.1; -;  
 DR EMBL; S81342; AAB36228.2; -;  
 DR EMBL; S81342; AAB36231.1; -;  
 DR EMBL; S81342; AAB36230.1; ALT\_SEQ.  
 DR HSP; P42224; IBF5.  
 DR MGD; MGI:103039; Stat2.  
 DR InterPro; IPR000980; SH2.  
 DR InterPro; IPR001217; STAT.  
 DR Pfam; PF00017; SH2; 1.  
 DR Pfam; PF01017; STAT; 1.  
 DR SMART; SM00252; SH2; 1.  
 DR PROSITE; PS00001; SH2; 1.  
 KW Transcription regulation; DNA-binding; Nuclear protein;  
 KW Phosphorylation; SH2 domain; Alternative splicing.  
 FT DOMAIN 571 666  
 FT MOD\_RES 689 689  
 FT VARSPLIC 620 643  
 FT HKVEIYSQPTPKVQLSLPLTEI -> GQHPPPVHSCSL  
 FT SARHPDRLPP (IN SHORT ISOFORM).  
 FT MISSING (IN SHORT ISOFORM).  
 FT T -> A (IN REF. 2).  
 FT CONFLICT 596 596  
 FT CONFLICT 620 620 H -> D (IN REF. 2).  
 SQ SEQUENCE 923 AA; 105416 MW; D50BB54C535B0774 CRC64;  
 -----  
 Query Match 52.5%; Score 32; DB 1; Length 923;  
 Best Local Similarity 85.7%; Pred. No. 2.8e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 -----  
 QY 2 RPKPQOFF 8  
 DB 478 RPKPQOFF 484  
 -----  
 RESULT 75  
 Y124.METJA  
 ID Y124.METJA STANDARD; PRT; 1075 AA.  
 AC Q57588;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE HYPOTHETICAL PROTEIN MJ0124.  
 GN MJ0124  
 OS Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
 OC Methanococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=96337999; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;

RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii";  
 RL Science 273:1058-1073(1996).  
 CC -!- SIMILARITY: TO M.JANNASCHII MJ0124.  
 CC -!- SIMILARITY: SOME, TO TYPE I RESTRICTION ENZYMES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U67469; AAB98104.1; -;  
 DR TIGR; MJ0124; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 1075 AA; 127796 MW; 4F765E19E0B52889 CRC64;  
 -----  
 Query Match 52.5%; Score 32; DB 1; Length 1075;  
 Best Local Similarity 54.5%; Pred. No. 3.3e+02;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 -----  
 QY 1 RPKPQOFFGLM 11  
 DB 690 RPYDPKEFGLI 700  
 -----  
 RESULT 76  
 RPOB.HETCA  
 ID RPOB.HETCA STANDARD; PRT; 1116 AA.  
 AC P36440;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE DNA-DIRECTED RNA POLYMERASE BETA CHAIN (EC 2.7.7.6).  
 GN RPOB.  
 OS Heterosigma carterae.  
 OG Chloroplast.  
 OC Eukaryota; Stramenopiles; Raphidophyceae; Heterosigma.  
 OX NCBI\_TaxID=28465;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Mangahas J.L., Cattolico R.A., Reynolds A.E.;  
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.  
 CC -!- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE = N PYROPHOSPHATE +  
 CC RNA(N).  
 CC -!- SUBUNIT: IN CHLOROPLAST THE RNA POLYMERASE IS COMPOSED OF FOUR  
 CC SUBUNITS: ALPHA, BETA, BETA', AND BETA".  
 CC -!- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X75815; CAA53450.1; -;  
 DR Mendel; 4095; HETCA; RPOB;1.  
 DR InterPro; IPR001572; RNA\_pol\_B.  
 DR Pfam; PF00562; RNA\_pol\_B; 1.  
 DR PROSITE; PS01166; RNA\_POL\_BETA; 1.  
 KW Transferase; DNA-directed RNA polymerase; Transcription; Chloroplast.  
 SQ SEQUENCE 1116 AA; 125818 MW; EC6C83C81234435B CRC64;  
 -----  
 Query Match 52.5%; Score 32; DB 1; Length 1116;

```
Best Local Similarity 50.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 RPKQOFFGL 10
DB 775 RPKPGKVFVG 784

Query Match 52.5%; Score 32; DB 1; Length 1206;
Best Local Similarity 71.4%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

SQ SEQUENCE 1206 AA; 133464 MW; 4DFB38CB52BD8EE7 CRC64;

OY 2 PKPQOFF 8
DB 1037 PEQDFF 1043

RESULT 77
FM14_MOUSE
ID FM14_MOUSE STANDARD; PRT; 1206 AA.
AC Q05859;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE FORMIN 1 ISOFORM IV (LIMB DEFORMITY PROTEIN).
GN FMN OR LD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=92112033; PubMed=1339380;
RA Grusby-Jackson L., Kuo A., Leder P.;
RT "A variant limb deformity transcript expressed in the embryonic mouse
limb defines a novel formin";
RL Genes Dev. 6:29-37(1992).
CC -1- FUNCTION: IS IMPORTANT IN THE MORPHOGENESIS OF LIMB AND MAY HAVE A
CC FUNCTION IN DIFFERENTIATED CELLS OR BE INVOLVED IN MAINTAINING
CC SPECIFIC DIFFERENTIATED STATES.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 5 ISOFORMS: IA (AC Q05860), IB (AC
CC Q05860), II (AC Q05860), III (AC Q05860) AND IV (SHOWN HERE); ARE
CC PRODUCED BY ALTERNATIVE SPLICING. A VARIATION IN SPLICING IS SEEN
CC AMONG DIFFERENT TISSUES AND DIFFERENT SIZE TRANSCRIPTS EXIST
CC WITHIN ANY ONE TISSUE.
CC -1- TISSUE SPECIFICITY: IT IS FOUND THROUGHOUT THE EMBRYO BUT
CC HAS A FUNCTIONAL ROLE ONLY IN THE KIDNEY AND LIMB.
CC -1- DEVELOPMENTAL STAGE: THIS IS THE ISOFORM FOUND IN THE APICAL
CC ECTODERMAL RIDGE AND THE MESENCHYMAL COMPARTMENT OF THE DEVELOPING
CC LIMB BUD.
CC -1- PTM: PHOSPHORYLATED ON SERINE AND POSSIBLY THREONINE RESIDUES.
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOMOLOGY 1 (FH1) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOMOLOGY 2 (FH2) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE FORMIN HOMOMOLOGY FAMILY. CAPPUCCINO
CC SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X62379; CAA44244.1; -.
CC PIR; S24407; S24407.
CC MGD; MGI:101815; Fmn.
CC InterPro; IPR003104; FH2.
CC InterPro; IPR001265; Formin.
CC Pfam; PF02181; FH2; 1.
CC PRINTS; PR00828; FORMIN.
CC SMART; SM00498; FH2; 1.
CC SMART; SM00498; FH2; 1.
CC Nuclear protein; Developmental protein; Alternative splicing;
CC Phosphorylation; Coiled coil.
CC DOMAIN 418 443 COILED COIL (POTENTIAL).
CC DOMAIN 497 566 COILED COIL (POTENTIAL).
CC DOMAIN 644 744 FH1 (PRO-RICH).
CC DOMAIN 759 1164 FH2.
CC DOMAIN 1043 1116 COILED COIL (POTENTIAL).
CC DOMAIN 635 638 POLY-SER.
CC DOMAIN 751 755 POLY-SER.
CC FT
CC FT

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; X62681; CAA44555.1; -.
InterPro; IPR003104; FH2.
InterPro; IPR001265; Formin.
Pfam; PF02181; FH2; 1.
PRINTS; PR00828; FORMIN.
SMART; SM00498; FH2; 1.
SMART; SM00498; FH2; 1.
Nuclear protein; Developmental protein; Coiled coil;
Alternative splicing.
DOMAIN 428 450 COILED COIL (POTENTIAL).
FT
FT
```

FT DOMAIN 503 572 COILED COIL (POTENTIAL).  
FT DOMAIN 652 751 FH1 (PRO-RICH).  
FT DOMAIN 766 1171 FH2.  
FT DOMAIN 1050 1125 COILED COIL (POTENTIAL).  
SQ SEQUENCE 1213 AA; 135240 MW; ADE3EF0B3FB9D862 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 1213;  
Best Local Similarity 71.4%; Pred. No. 3.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQOFF 8  
I:|I|I|  
Db 1044 PEQDFF 1050

## RESULT 79

FMN1\_MOUSE STANDARD; PRT; 1468 AA.  
AC Q05860;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE FORMIN 1 ISOFORMS I/II/III (LIMB DEFORMITY PROTEIN).  
GN FMN OR LD.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney, and Testis;  
RX MEDLINE=90363291; PubMed=2392150;  
RA Woychik R.P., Maas R.L., Zeller R., Vogt T.F., Leder P.;  
RT "Formins": proteins deduced from the alternative transcripts of the  
RT limb deformity gene.";  
RL Nature 346:850-853(1990).  
RN [2]  
RN ALTERNATIVE SPLICING  
RX MEDLINE=97224459; PubMed=9119367;  
RA Wang C.C., Chan D.C., Leder P.;  
RT "The mouse formin (Fmn) gene: genomic structure, novel exons, and  
RT genetic mapping.";  
RL Genomics 39:303-311(1997).  
RN [3]  
RP PHOSPHORYLATION.  
RX MEDLINE=93296176; PubMed=8516300;  
RA Vogt T.F., Jackson-Grusby L., Rush J., Leder P.;  
RT "Formins: phosphoprotein isoforms encoded by the mouse limb deformity  
RT locus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:5554-5558(1993).  
CC -!- FUNCTION: IS IMPORTANT FOR THE MORPHOGENESIS OF LIMB AND KIDNEY  
CC AND MAY HAVE A FUNCTION IN DIFFERENTIATED CELLS OR MAY BE  
CC INVOLVED IN MAINTAINING SPECIFIC DIFFERENTIATED STATES.  
CC -!- SUBCELLULAR LOCATION: NUCLEAR.  
CC -!- ALTERNATIVE PRODUCTS: AT LEAST 5 ISOFORMS: IA (SHOWN HERE), IB,  
CC II, III AND IV (AC Q05860). ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC A VARIATION IN SPLICING IS SEEN AMONG DIFFERENT TISSUES AND  
CC DIFFERENT SIZE TRANSCRIPTS EXIST WITHIN ANY ONE TISSUE.  
CC -!- TISSUE SPECIFICITY: IT IS PRESENT IN THE ADULT KIDNEY, TESTIS,  
CC LIMB, OVARY, BRAIN, SMALL INTESTINE, SALIVARY GLAND AND HARDERIAN  
CC GLAND. IT IS PRESENT THROUGHOUT THE EMBRYO.  
CC -!- DEVELOPMENTAL STAGE: IN THE DEVELOPING LIMB BUD, THE PROTEIN  
CC IS EXPRESSED IN THE APICAL ECTODERMAL RIDGE AND THE MESENCHYMAL  
CC COMPARTMENT, PREDOMINANTLY IN THE POSTERIOR REGION. DURING  
CC KIDNEY MORPHOGENESIS, EXPRESSION IS INITIALLY RESTRICTED TO  
CC THE EPITHELIAL COMPARTMENT OF THE PRONEPHROS AND MESONEPHROS.  
CC -!- PTM: PHOSPHORYLATED ON SERINE AND POSSIBLY THREONINE RESIDUES.  
CC -!- SIMILARITY: CONTAINS 1 FORMIN HOMOLGY 1 (FH1) DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 FORMIN HOMOLGY 2 (FH2) DOMAIN.  
CC -!- SIMILARITY: BELONGS TO THE FORMIN HOMOLGY FAMILY. CAPPUPCINO  
CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; X53599; CAA37668.1; -  
DR PIR; S11515; S11515.  
DR MGD; MGI:101815; Fmn.  
DR InterPro; IPR003104; FH2.  
DR Pfam; PF02181; FH2; 1.  
DR PRINTS; PR00828; FORMIN.  
DR SMART; SM00498; FH2; 1.  
KW Nuclear protein; Developmental protein; Alternative splicing;  
KW Phosphorylation; Coiled coil.  
FT DOMAIN 723 792 COILED COIL (POTENTIAL).  
FT DOMAIN 870 970 FH1 (PRO-RICH).  
FT DOMAIN 985 1426 FH2.  
FT DOMAIN 1305 1378 COILED COIL (POTENTIAL).  
FT DOMAIN 198 203 POLY-SER.  
FT DOMAIN 861 864 POLY-SER.  
FT DOMAIN 885 892 POLY-PRO.  
FT DOMAIN 911 925 POLY-PRO.  
FT DOMAIN 929 940 POLY-PRO.  
FT DOMAIN 951 962 POLY-PRO.  
FT DOMAIN 966 970 POLY-PRO.  
FT DOMAIN 977 981 POLY-SER.  
FT VARSPLIC 1252 1287 MISSING (IN ISOFORM IB).  
FT VARSPLIC 625 722 MISSING (IN ISOFORM II).  
FT VARSPLIC 626 627 IA -> SV (IN ISOFORM III).  
FT VARSPLIC 628 1468 MISSING (IN ISOFORM III).  
SQ SEQUENCE 1468 AA; 163809 MW; EF2FB1E9CA9DAF43 CRC64;

Query Match 52.5%; Score 32; DB 1; Length 1468;  
Best Local Similarity 71.4%; Pred. No. 4.4e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 PKPQOFF 8  
I:|I|I|  
Db 1299 PEQDFF 1305

## RESULT 80

VIT6\_CAEEL STANDARD; PRT; 1651 AA.  
AC P18948;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE VITELLOGENIN 6 PRECURSOR.  
GN VIT-6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=91251142; PubMed=1904098;  
RA Spieth J., Nettleton M., Zucker-Aprison E., Lea K., Blumenthal T.;  
RT "Vitellogenin motifs conserved in nematodes and vertebrates.";  
RL J. Mol. Evol. 32:429-438(1991).  
RN [2]  
RP SEQUENCE OF 1-110 FROM N.A.  
RX MEDLINE=86284606; PubMed=3841791;  
RA Spieth J., Blumenthal T.;  
RT "The Caenorhabditis elegans vitellogenin gene family includes a gene  
RT encoding a distantly related protein.";  
RL Mol. Cell. Biol. 5:2495-2501(1985).

```

RN SEQUENCE OF 1-24 FROM N.A.
RX MEDLINE=85269643; PubMed=4022780;
RA Splith J., Denison K., Kirtland S., Cane J., Blumenthal T.;
RT "The C. elegans vitellogenin genes: short sequence repeats in the
RL Nucleic Acids Res. 13:5283-5295(1985).
CC -1- FUNCTION: PRECURSOR OF THE EGG-YOLK PROTEINS THAT ARE SOURCES OF
CC NUTRIENTS DURING EMBRYONIC DEVELOPMENT (POTENTIAL).
CC -1- TISSUE SPECIFICITY: SYNTHESIZED IN CAENORHABDITIS ONLY BY 32 CELLS
CC BUILDING THE INTESTINE OF ADULT HERMAPHRODITIC INDIVIDUALS; THEY
CC ARE COTRANSLATIONALLY SECRETED INTO THE BODY CAVITY & SUBSEQUENTLY
CC TAKEN UP BY THE GONAD.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56213; CAA39670.1; -
CC EMBL; M11499; AAA28165.1; -
CC PIR; A27271; A27271.
CC InterPro: IPR001747; Vitellogenin_N.
CC InterPro: IPR001846; Vwd.
CC Pfam; PF01347; Vitellogenin_N; 1.
CC Pfam; PF00094; vwd; 1.
CC SMART; SM00216; VWD; 1.
CC Storage protein; Multigene family; Signal.
CC SIGNAL 1 15 POTENTIAL.
CC CHAIN 16 1651 VITELLOGENIN 6.
CC SEQUENCE 1651 AA; 193400 MW; 2CDZF72C3EA374F7 CRC64;
CC -----
Query Match 52.5%; Score 32; DB 1; Length 1651;
Best Local Similarity 71.4%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 2 PKPQFF 8
DB 1234 PKPAQYF 1240
-----
RESULT 81
ID SC16_YEAST STANDARD; PRT; 2194 AA.
AC P48415;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE MULTIDOMAIN VESICLE COAT PROTEIN.
GN SC16 OR YPL085W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96017704; PubMed=7593161;
RA Espenshade P., Gimeno R.E., Holzmacher E., Teung P., Kaiser C.A.;
RT "Yeast Sc16 gene encodes a multidomain vesicle coat protein that
RT interacts with Sec23p".
RL J. Cell Biol. 131:311-324(1995).
CC -1- FUNCTION: INVOLVED IN THE BUDDING OF TRANSPORT VESICLE FROM THE
CC ENDOPLASMIC RETICULUM. THE C-TERMINAL INTERACTS WITH SEC23 AND
CC WITH THE CYTOSOLIC DOMAIN OF SED4. COULD THEREFORE BE A
CC CONSTITUENT OF COPII VESICLE COAT. N-TERMINAL OVEREXPRESSION
CC CAUSES A LETHAL SECRETION DEFECT.
CC -1- SUBCELLULAR LOCATION: ON THE ENDOPLASMIC RETICULUM AND ON VESICLES
CC WHICH BUD FROM IT.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U23819; AAC49088.1; -
CC SGD; S0006006; SEC16.
CC Transport; Protein transport; Golgi stack; Endoplasmic reticulum.
CC L-S: IN SEC16-4; TS ACCUMULATION OF ER
CC MUTAGEN 1058 1058 MEMBRANES.
CC -----
CC MUTAGEN 1083 1083 L-S: IN SEC16-3; TS ACCUMULATION OF ER
CC MEMBRANES.
CC MUTAGEN 1088 1088 L-S: IN SEC16-2; TS ACCUMULATION OF ER
CC MEMBRANES.
CC MUTAGEN 1230 1230 W-R: IN SEC16-1; TS ACCUMULATION OF ER
CC MEMBRANES.
CC SEQUENCE 2194 AA; 241613 MW; BB1E02D2AD4683E3 CRC64;
CC -----
Query Match 52.5%; Score 32; DB 1; Length 2194;
Best Local Similarity 85.7%; Pred. No. 6.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4 PQQFFGL 10
DB 581 PQQFHL 587
-----
RESULT 82
ID TRN2_UPERU STANDARD; PRT; 11 AA.
AC P08616;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE RUGOSAUPEROLEIN II (LVS(5)-THR(6)-PHYSALAEMIN).
OS Uperoleia rugosa (Australian leptodactylid frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Myobatrachidae;
OC Uperoleia.
OX NCBI_TaxID=8368;
RN [1]
RP SEQUENCE.
RX MEDLINE=80223080; PubMed=7389029;
RA Nakajima T., Yasuhara T., Erspamer V., Erspamer G.F., Negri L.;
RT "Physalaemin- and bombesin-like peptides in the skin of the
RT Australian leptodactylid frog Uperoleia rugosa.";
RL Chem. Pharm. Bull. 28:689-695(1980).
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
CC InterPro: IPR003580; Protachykinin.
CC InterPro: IPR002040; Tachykinin.
CC Pfam; PF02202; Tachykinin; 1.
CC SMART; SM00203; TK; 1.
CC PROSITE; PS00267; TACHYKININ; 1.
CC Tachykinin; Neuropeptide; Amidation; Amphibian skin.
CC MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
CC MOD_RES 11 11 AMIDATION.
CC SEQUENCE 11 AA; 1270 MW; 3293693E59D1A327 CRC64;
CC -----
Query Match 50.8%; Score 31; DB 1; Length 11;
Best Local Similarity 62.5%; Pred. No. 5.1;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 4 PQQFFGLM 11

```

```
Db 4 PRFTYGLM 11
|: |::|
|: |::|

RESULT 83
MOT2_MERUN STANDARD; PRT; 71 AA.
ID MOT2_MERUN AC O35440;
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MONOCARBOXYLATE TRANSPORTER 2 (MCT 2) (FRAGMENT).
GN SLC16A7 OR MCT2.
OS Meriones unguiculatus (Mongolian jird).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC Meriones.
OX NCBI_TaxID=10047;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Stria vascularis;
RX MEDLINE=98107623; PubMed=9447934;
RA Shinozono M., Scofield M.A., Wangemann P.;
RT "Functional evidence for a monocarboxylate transporter (MCT) in stria
RT marginal cells and molecular evidence for MCT1 and MCT2 in stria
RT vascularis."
RL Hear. Res. 114:213-222(1997).
CC -!- FUNCTION: PROTON-LINKED MONOCARBOXYLATE TRANSPORTER. CATALYZES THE
CC RAPID TRANSPORT ACROSS THE PLASMA MEMBRANE OF MANY
CC MONOCARBOXYLATES SUCH AS LACTATE, PYRUVATE, BRANCHED-CHAIN OXO
CC ACIDS DERIVED FROM LEUCINE, VALINE AND ISOLEUCINE, AND THE KETONE
CC BODIES ACETOACETATE, BETA-HYDROXYBUTYRATE AND ACETATE (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PLASMA MEMBRANE
CC (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE SLC16 FAMILY OF TRANSPORTERS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; AF029767; AAB84219.1;
FT NON_TER 1
FT TRANSMEM <1 14 POTENTIAL.
FT TRANSMEM 24 44 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT NON_TER 71 71
SQ SEQUENCE 71 AA; 7864 MW; 25C82B27F22B61B1 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 71;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQFFGL 10
|: |::|
|: |::|
Db 21 RPRIQYFFSL 30

RESULT 84
FTRV_MAIZE
ID FTRV_MAIZE STANDARD; PRT; 97 AA.
AC P80680;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE FERREDOXIN-THIOREDOXIN REDUCTASE, VARIABLE CHAIN (FTR-V) (FERREDOXIN-
DE THIOREDOXIN REDUCTASE SUBUNIT A) (FTR-A).

Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE.
RC TISSUE=Leaf;
RX MEDLINE=97054599; PubMed=8898896;
RA Iwadata H., Tsugita A., Chow L.-P., Kizuki K., Stritt-Etter A.-L.,
RA Li J., Schuermann P.;
RT "Amino acid sequence of the maize ferredoxin:thioredoxin reductase
RT variable subunit."
RL Eur. J. Biochem. 241:121-125(1996).
CC -!- FUNCTION: FTR IS A [4FE-4S] PROTEIN PLAYING A CENTRAL ROLE IN THE
CC FERREDOXIN/THIOREDOXIN REGULATORY CHAIN. IT CONVERTS AN ELECTRON
CC SIGNAL (PHOTOREDOXED FERREDOXIN) TO A THIOL SIGNAL (REDUCED
CC THIOREDOXIN) IN THE REGULATION OF ENZYMES BY REDUCTION OF SPECIFIC
CC DISULFIDE GROUPS. CATALYZES THE LIGHT-DEPENDENT ACTIVATION OF
CC SEVERAL PHOTOSYNTHETIC ENZYMES.
CC -!- SUBUNIT: HETERODIMER OF SUBUNIT A (VARIABLE SUBUNIT) AND SUBUNIT
CC B (CATALYTIC SUBUNIT).
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC -!- SIMILARITY: TO SPINACH AND SYNECHOCOCCUS SP. FTR-V.
DR MaizedB; 134030;
KW Oxidoreductase; Chloroplast.
SQ SEQUENCE 97 AA; 10886 MW; DE1ED2AEE76B0FF5 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 97;
Best Local Similarity 62.5%; Pred. No. 45;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFF 8
|: |::|
|: |::|
Db 77 QPKPVRFF 84

RESULT 85
YD43_MYCLE
ID YD43_MYCLE STANDARD; PRT; 126 AA.
AC P54134;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE HYPOTHETICAL 14.5 KDA PROTEIN ML1177.
GN ML1177 OR MLCB1701.03C OR B1549_F3_106.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith D.R., Robison K.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagers K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.-A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -!- SIMILARITY: STRONG, TO M.TUBERCULOSIS RV1343C.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U00014; AAA50903.1; -;  
DR EMBL; AL049191; CAB39143.1; -;  
DR EMBL; AL583921; CAC31558.1; -;  
DR Leproma; MLI177; -;  
DR PROSITE; PS00013; PROKAR\_LIPOPROTEIN; UNKNOWN\_1.  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 10 30 POTENTIAL.  
FT TRANSMEM 44 64 POTENTIAL.  
SQ SEQUENCE 126 AA; 14479 MW; 687096956561908D CRC64;

Query Match 50.8%; Score 31; DB 1; Length 126;  
Best Local Similarity 53.3%; Pred. No. 59;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RPKPQQ 6  
||||:|  
Db 90 RPKPEQ 95

## RESULT 86

RS16\_YEAST  
ID RS16\_YEAST STANDARD; PRT; 142 AA.  
AC P40213; P26787;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE 40S RIBOSOMAL PROTEIN S16 (RP61R).  
GN (RPS16A OR RP61R OR YMR143W OR YN9375.12) AND (RPS16B OR YDL083C).  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A. (RPS16A).  
RC STRAIN=S288C / AB972;  
RA Badcock K., Churcher C., Barrell B.G., Rajandream M.A., Walsh S.V.;  
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.  
RT [2]  
RP SEQUENCE FROM N.A. (RPS16B).  
RA Wambutt R., Wedler H., Wedler E., Scharfe M.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RT [3]  
RP PRELIMINARY PARTIAL SEQUENCE OF 1-25.  
RX MEDLINE=92184799; PubMed=1544921;  
RA Takakura H., Tsunasawa S., Miyagi M., Warner J.R.;  
RT "NH2-terminal acetylation of ribosomal proteins of Saccharomyces  
RT cerevisiae";  
RL J. Biol. Chem. 267:5442-5445(1992).  
CC -!- MISCELLANEOUS: THERE ARE TWO GENES FOR S16 IN YEAST.  
CC -!- SIMILARITY: BELONGS TO THE S9P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; Z47071; CAA87357.1; -;  
DR EMBL; Z74131; CAA98649.1; -;  
DR SGD; S0004751; RPS16A.  
DR SGD; S0002241; RPS16B.  
DR InterPro; IPR000754; Ribosomal\_S9.  
DR Pfam; PF00380; Ribosomal\_S9; 1.

DR PRODOM; PD001627; Ribosomal\_S9; 1.  
DR PROSITE; PS00360; RIBOSOMAL\_S9; 1.  
KW Ribosomal protein; Acetylation; Multigene family.  
FT INIT\_MET 0 0  
FT MOD\_RES 1 1 ACETYLTATION.  
SQ SEQUENCE 142 AA; 15716 MW; 15873374B3262144 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 142;  
Best Local Similarity 55.6%; Pred. No. 66;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQFFG 9  
||||:|  
Db 122 RPEPKKEGG 130

## RESULT 87

PR39\_PIG  
ID PR39\_PIG STANDARD; PRT; 172 AA.  
AC P80054; Q9TR84;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ANTIBACTERIAL PROTEIN PR-39 PRECURSOR.  
GN PR39.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95350216; PubMed=7624374;  
RA Gudmundsson G.H., Magnusson K.P., Chowdhary B.P., Johansson M.,  
RA Andersson L., Boman H.G.;  
RT "Structure of the gene for porcine peptide antibiotic PR-39, a  
RT cathelin gene family member: comparative mapping of the locus for the  
RT human peptide antibiotic FALL-39";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:7085-7089(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Bone marrow;  
RX MEDLINE=94071853; PubMed=8250863;  
RA Storici P., Zanetti M.;  
RT "A cDNA derived from pig bone marrow cells predicts a sequence  
RT identical to the intestinal antibacterial peptide PR-39";  
RL Biochem. Biophys. Res. Commun. 196:1058-1065(1993).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=96105365; PubMed=7498526;  
RA Zhao C., Ganz T., Lehrer R.I.;  
RT "Structures of genes for two cathelin-associated antimicrobial  
RT peptides: prophenin-2 and PR-39";  
RL FEBS Lett. 376:130-134(1995).  
RN [4]  
RP SEQUENCE OF 131-169.  
RC TISSUE=Intestine;  
RX MEDLINE=92111534; PubMed=1765098;  
RA Agerberth B., Lee J.-Y., Bergman T., Carlquist M., Boman H.G.,  
RA Mutt V., Joernvall H.;  
RT "Amino acid sequence of PR-39. Isolation from pig intestine of a new  
RT member of the family of proline-arginine-rich antibacterial  
RT peptides";  
RL Eur. J. Biochem. 202:849-854(1991).  
RN [5]  
RP SEQUENCE OF 131-164, AND FUNCTION.  
RC TISSUE=Neutrophils;  
RX MEDLINE=95088504; PubMed=7996056;  
RA Shi J., Ross C.R., Chengappa M.M., Blecha P.;  
RT "Identification of a proline-arginine-rich antibacterial peptide from  
RT neutrophils that is analogous to PR-39, an antibacterial peptide from  
RT the small intestine";



RL J. Leukoc. Biol. 56:807-811(1994).  
CC -1- FUNCTION: EXERTS A POTENT ANTIMICROBIAL ACTIVITY AGAINST BOTH  
CC E.COLI AND B.MEGATERIUM.  
CC -1- TISSUE SPECIFICITY: SMALL INTESTINE AND BONE MARROW.  
CC -1- SIMILARITY: BELONGS TO THE CATHelicidin FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X87236; CAA60682.1; -  
CC EMBL: L23825; AAA31109.1; -  
CC EMBL: X89201; CAA61487.1; -  
CC PIR: S19563; S19563.  
CC InterPro: IPR001894; Cathelicidin.  
CC Pfam: PF00666; Cathelicidins; 1.  
CC ProDom: PD001838; Cathelicidin; 1.  
CC PROSITE: PS00946; CATHelicidins\_1; 1.  
CC PROSITE: PS00947; CATHelicidins\_2; 1.  
KW Antibiotic; Amidation; Signal.  
FT SIGNAL 1 29  
FT PROPEP 30 130  
FT CHAIN 131 169  
FT MOD\_RES 30 30  
FT -----  
FT FT DISULFID 85 96  
FT FT DISULFID 107 124  
FT FT MOD\_RES 169 169  
FT FT CONFLICT 21 21  
FT FT CONFLICT 29 29  
FT FT CONFLICT 90 91  
FT FT CONFLICT 117 119  
FT FT CONFLICT 157 157  
SQ SEQUENCE 172 AA; 19476 MW; 994B792798C0E133 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 172;  
Best Local Similarity 62.5%; Pred. No. 80;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQFF 8  
DB 141 RRPDPFF 148  
||:|||||

RESULT 88  
TBP\_ARCFU STANDARD; PRT; 183 AA.  
ID TBP\_ARCFU STANDARD; PRT; 183 AA.  
AC O29874;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE TATA-BOX BINDING PROTEIN (TATA-BOX FACTOR) (TATA SEQUENCE-BINDING  
DE PROTEIN) (TBP) (BOX A BINDING PROTEIN) (BAP).  
GN TBP OR AF0373.  
OS Archaeoglobus fulgidus.  
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;  
OC Archaeoglobus.  
OX NCBI\_TaxID=2234;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;  
RX MEDLINE=98049343; PubMed=9389475;  
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,  
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,  
RA Richardson D.L., Kurlavich A.R., Graham D.E., Kyprides N.C.,  
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,  
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,  
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,

RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,  
RA Cotton M.D., Spriggs T., Artiaga P., Kaine B.P., Sykes S.M.,  
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,  
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,  
RA Venter J.C.;  
RT "The complete genome sequence of the hyperthermophilic, sulphate-  
RT reducing archaeon Archaeoglobus fulgidus.";  
RL Nature 390:364-370(1997).  
CC -1- FUNCTION: GENERAL FACTOR THAT PLAYS A ROLE IN THE ACTIVATION OF  
CC ARCHAEL GENES TRANSCRIBED BY RNA POLYMERASE. BINDS SPECIFICALLY  
CC TO THE TATA BOX PROMOTER ELEMENT WHICH LIES CLOSE TO THE POSITION  
CC OF TRANSCRIPTION INITIATION (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE TBP FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AE001078; AAB90862.1; -  
CC TIGR: AF0373; -  
CC InterPro: IPR000814; TFIID.  
CC Pfam: PF00352; TBP; 2.  
CC PRINTS: PR00686; TIFACTORIID.  
CC PROSITE: PS00351; TFIID; 1.  
KW Transcription regulation; DNA-binding; Repeat; Complete proteome.  
FT REPEAT 8 84  
FT REPEAT 99 177  
SQ SEQUENCE 183 AA; 20135 MW; 255C5C05B8E3D8A3 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 183;  
Best Local Similarity 66.7%; Pred. No. 85;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFFGLM 11  
DB 36 RPKQPPGLV 44  
||:|||||

RESULT 89  
MOVP\_TOML STANDARD; PRT; 264 AA.  
ID MOVP\_TOML STANDARD; PRT; 264 AA.  
AC P03584;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE MOVEMENT PROTEIN (CELL-TO-CELL TRANSPORT PROTEIN) (30 KDA PROTEIN).  
GN MP.  
OS Tomato mosaic virus (strain L) (TOMV) (TMV strain tomato),  
OS Tomato mosaic virus (strain Kazakh K1) (TOMV) (TMV strain K1), and  
OS Tomato mosaic virus (strain Kazakh K2) (TOMV) (TMV strain K2).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.  
OX NCBI\_TaxID=12252, 138311, 138312;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=L;  
RX MEDLINE=83220776; PubMed=6304642;  
RA Takamatsu N., Ohno T., Meshi T., Okada Y.;  
RT "Molecular cloning and nucleotide sequence of the 30K and the coat  
RT protein cistron of TMV (tomato strain) genome.";  
RL Nucleic Acids Res. 11:3767-3778(1983).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=L;  
RX MEDLINE=85157522; PubMed=6549393;  
RA Ohno T., Aoyagi M., Yamanashi Y., Saito H., Ikawa S., Meshi T.,  
RA Okada Y.;  
RT "Nucleotide sequence of the tobacco mosaic virus (tomato strain)  
RT genome and comparison with the common strain genome.";



Query Match 50.8%; Score 31; DB 1; Length 264;  
Best Local Similarity 71.48; Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQF 7  
||||:|  
DB 233 RPKPKSF 239

## RESULT 92

ID YOTB\_CAEEL STANDARD; PRT; 266 AA.  
AC P34657;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL 30.2 KDA PROTEIN ZK632.12 IN CHROMOSOME III.  
GN ZK632.12.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=94150718; PubMed=7906398;  
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M., Coulson A.,  
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Fraser A.,  
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,  
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,  
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,  
RA Sims M., Smaldon N., Smith A., Smith K., Vaudin M., Vaughan K.,  
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
RA Watson R., Watson A., Weinstock L., Wilkinson-Sproat J.,  
RA Wohlman P.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
RT elegans";  
RL Nature 368:32-38(1994).  
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 FYVE-TYPE ZINC FINGER.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; Z22181; CAA80187.1; -;  
WormPep; ZK632.12; CE01110.  
InterPro; IPR001849; PH.  
InterPro; IPR000306; Znf\_FYVE.  
Pfam; PF01363; FYVE; 1.  
Pfam; PF00169; PH; 1.  
SMART; SM00064; FYVE; 1.  
SMART; SM00233; PH; 1.  
PROSITE; PS50178; ZF\_FYVE; 1.  
PROSITE; PS50003; PH\_DOMAIN; 1.  
KW Hypothetical protein: Zinc-finger.  
FT DOMAIN 35 131 PH.  
FT ZN\_FING 152 212 FYVE-TYPE.  
SQ SEQUENCE 266 AA; 30187 MW; 91C2F62EDF13839E CRC64;

Query Match 50.8%; Score 31; DB 1; Length 266;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFF 8

DB 49 KPKQKQFF 56  
|||:|

## RESULT 93

ID THIM\_PASMU STANDARD; PRT; 267 AA.  
AC P57931;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYDROXYETHYLTHIAZOLE KINASE (EC 2.7.1.50) (4-METHYL-5-BETA-  
DE HYDROXYETHYLTHIAZOLE KINASE) (THZ KINASE) (TH KINASE).  
GN THIM OR PML262.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PM70;  
RX MEDLINE=21145866; PubMed=11248100;  
RA May B.J., Zhang Q., Li L.L., Faustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -1- CATALYTIC ACTIVITY: ATP + 4-METHYL-5-(2-HYDROXYETHYL)-THIAZOLE =  
CC ADP + 4-METHYL-5-(2-PHOSPHOETHYL)-THIAZOLE.  
CC -1- COFACTOR: MAGNESIUM (BY SIMILARITY).  
CC -1- PATHWAY: THIAMINE BIOSYNTHESIS.  
CC -1- SIMILARITY: BELONGS TO THE THZ KINASE FAMILY.  
CC -----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; AE006165; AAK03346.1; -;  
InterPro; IPR000417; Hyethyz\_kinase.  
Pfam; PF02110; HK; 1.  
KW Thiamine biosynthesis; Transferase; Kinase; ATP-binding; Magnesium;  
Complete proteome.  
FT METAL 91 91 MAGNESIUM (BY SIMILARITY).  
FT METAL 123 123 MAGNESIUM (BY SIMILARITY).  
FT ACT\_SITE 194 194 BASE (BY SIMILARITY).  
SQ SEQUENCE 267 AA; 28245 MW; C710B90C6BB5E971 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 267;  
Best Local Similarity 55.6%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPKQFFGLM 11  
|||:|  
DB 208 KPEQYFDM 216

## RESULT 94

ID CCHL\_YEAST STANDARD; PRT; 269 AA.  
AC P06182;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE CYTOCHROME C HEME LYASE (EC 4.4.1.17) (CCHL) (HOLOGYCYTOCHROME-C  
DE SYNTHASE).  
GN CYC3 OR YAL039C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;

```

RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-B-7034;
RX  MEDLINE=87218469; PubMed=3034577;
RA  Dumont M.E., Ernst J.F., Hampsey D.M., Sherman F.;
RT  "Identification and sequence of the gene encoding cytochrome c heme
RL  lyase in the yeast Saccharomyces cerevisiae.";
RL  EMBO J. 6:235-241(1987).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-S288C / AB972;
RX  MEDLINE=95249563; PubMed=7731988;
RA  Bussey H., Kaback D.B., Zhong W., Vo D.T., Clark M.W., Fortin N.;
RT  Hall J., Ouellette B.F.F., Keng T., Barton A.B., Su Y., Davies C.K.,
RA  Storms R.K.;
RT  "The nucleotide sequence of chromosome I from Saccharomyces
RT  cerevisiae.";
RL  Proc. Natl. Acad. Sci. U.S.A. 92:3809-3813(1995).
CC  -1- FUNCTION: LINKS COVALENTLY THE HEME GROUP TO THE APOPROTEIN
CC  OF CYTOCHROME C.
CC  -1- CATALYTIC ACTIVITY: HOLOCYTOCHROME C = APOCYTOCHROME C + HEME.
CC  -1- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC  -1- SIMILARITY: BELONGS TO THE CYTOCHROME C-TYPE HEME LYASE FAMILY.
CC  -1- SIMILARITY: CONTAINS 2 HEME REGULATORY MOTIFS (HRM).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
EMBL; X04776; CAA28470.1; -
DR  EMBL; U12980; AAC04992.1; -
DR  PIR; A26162; A26162.
DR  SGD; S0000037; CYC3.
DR  InterPro; IPR000511; Cyto_heme_lyase.
DR  Pfam; PF01265; Cyto_heme_lyase; 1.
DR  PROSITE; PS00821; CYTO_HEME_LYASE_1; 1.
DR  PROSITE; PS00822; CYTO_HEME_LYASE_2; 1.
KW  Lyase; Heme; Mitochondrion; Repeat.
FT  DOMAIN 25 30 HRM 1 (POTENTIAL).
FT  DOMAIN 41 46 HRM 2 (POTENTIAL).
SQ  SEQUENCE 269 AA; 30081 MW; A672A48BBD848AF CRC64;

Query Match 50.8%; Score 31; DB 1; Length 269;
Best Local Similarity 50.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PKPOFFGLM 11
Db 95 PPOQMYNAM 104

RESULT 95
RRP1_YEAST
ID RRP1_YEAST STANDARD; PRT; 278 AA.
AC P35178;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE RIBOSOMAL RNA PROCESSING PROTEIN 1.
GN RRP1 OR YDR087C OR D4478.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C / FY1679;
RA Coster F., Jonniaux J.-L., Goffeau A.;
```

```

RL Submitted (OCT-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 146-278 FROM N.A.
RX MEDLINE=94038890; PubMed=8223425;
RA Esnault Y., Blondel M.-O., Deshaies R.J., Schekman R., Kepes F.;
RT "The yeast S5S1 gene is essential for secretory protein translocation
RT and encodes a conserved protein of the endoplasmic reticulum.";
RL EMBO J. 12:4083-4093(1993).
RN [3]
RP CHARACTERIZATION.
RX MEDLINE=99276573; PubMed=10341208;
RA Savino T.M., Bastos R., Jansen E., Hernandez-Verdun D.;
RT "The nucleolar antigen Nop52, the human homologue of the yeast
RT ribosomal RNA processing RRP1, is recruited at late stages of
RT nucleogenesis.";
RL J. Cell Sci. 112:1889-1900(1999).
CC -1- FUNCTION: REQUIRED FOR 27S RNA PROCESSING TO 25S AND 5.8S.
CC -1- SUBCELLULAR LOCATION: NUCLEAR; NUCLEOLUS.
CC -1- SIMILARITY: BELONGS TO THE NRP-1 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; Z46796; CAA86809.1; -
DR  EMBL; Z74383; CAA98907.1; -
DR  EMBL; X82086; CAA57616.1; -
DR  EMBL; X74499; CAA52607.1; -
DR  PIR; S48776; S48776.
DR  SGD; S0002494; RRP1.
DR  Nuclear protein; rRNA processing; Coiled coil.
FT  DOMAIN 266 274 POLY-GLU.
SQ  SEQUENCE 278 AA; 33202 MW; 7E906A028ADA8A6 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 278;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQ 6
Db 58 RRPQQ 63

RESULT 96
NUGM_NEUCR
ID NUGM_NEUCR STANDARD; PRT; 283 AA.
AC P23710;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE 30.4 KDA SUBUNIT, MITOCHONDRIAL
DE PRECURSOR (EC 1.6.5.3) (EC 1.6.99.3) (COMPLEX I-30KD) (CI-31KD).
GN NUO-31.
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91024977; PubMed=2145832;
RA Videla A., Tropschug W., Werner S.;
RT "Primary structure and expression of a nuclear-coded subunit of
RT complex I homologous to proteins specified by the chloroplast
RT genome.";
RL Biochem. Biophys. Res. Commun. 171:1168-1174(1990).
CC -1- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE.
```

CC -!- CATALYTIC ACTIVITY: NADH + UBIQUINONE = NAD(+) + UBIQUINOL.  
CC -!- SUBUNIT: COMPLEX I IS COMPOSED OF ABOUT 30 DIFFERENT SUBUNITS.  
CC THIS IS A COMPONENT OF THE IRON-SULFUR PROTEIN FRACTION.  
CC -!- SUBCELLULAR LOCATION: MATRIX AND CYTOPLASMIC SIDE OF THE  
CC MITOCHONDRIAL INNER MEMBRANE.  
CC -!- SIMILARITY: BELONGS TO THE COMPLEX I 30 KDA SUBUNIT FAMILY.  
DR PR; A35935; A35935.  
DR InterPro: IPR001268; Complex1\_30K.  
DR Pfam: PF00329; complex1\_30Kd; 1.  
DR ProDom: PD001581; Complex1\_30K; 1.  
DR PROSITE; PS00542; COMPLEX1\_30K; 1.  
KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transit peptide.  
FT TRANSIT 1 17 MITOCHONDRION (POTENTIAL).  
FT CHAIN 18 283 NADH-UBIQUINONE OXIDOREDUCTASE 30.4 KDA  
FT SUBUNIT.  
SQ SEQUENCE 283 AA; 32283 MW; 3A2DCD32535986CA CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 283;  
Best Local Similarity 71.4%; Pred. No. 1.3e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPKPQQF 7  
DB 40 RPNRQF 46  
II I I I  
II I I I  
  
RESULT 97  
LST\_HAEIN STANDARD; PRT; 304 AA.  
ID LST\_HAEIN  
AC Q48211; Q05084;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE CMP-N-ACETYLNEURAMINATE-BETA-GALACTOSAMIDE-ALPHA-2,3-SIALYLTRANSFERASE  
DE (EC 2.4.99.-) (BETA-GALACTOSIDE ALPHA-2,3-SIALYLTRANSFERASE) (ALPHA  
DE 2,3-ST) (LIPOPOLISACCHARIDE SIALYLTRANSFERASE).  
GN LST OR H11699.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OX NCBI\_TaxID=727;  
[1]  
SEQUENCE FROM N.A.  
RA McLaughlin R., Abu Kwaik Y., Young R., Spinola S., Apicella M.;  
RT "Characterization and sequence of the *lsf* locus from *Haemophilus*  
RT *influenzae*.";  
RL Submitted (JUN-1992) to the EMBL/GenBank/DBJ databases.  
[2]  
SEQUENCE FROM N.A.  
RA STRAIN-RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Uterback T.R., Hanna M.C., Spriggs T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*  
RT Rd";  
RL Science 269:496-512(1995).  
CC -!- FUNCTION: TRANSFERS SIALIC ACID FROM THE SUBSTRATE CMP-SIALIC ACID  
CC DONOR TO THE TERMINAL BETA-D-GALACTOSYL-1,4-ACETYL-BETA-D-  
CC LIPOOLIGOSACCHARIDE.  
CC -!- CATALYTIC ACTIVITY: CMP-N-ACETYLNEURAMINATE + BETA-D-GALACTOSYL-  
CC 1,4-ACETYL-BETA-D-GLUCOSAMINE = CMP + ALPHA-N-ACETYLNEURAMINYL-  
CC 2,3-BETA-D-GALACTOSYL-1,4-N-ACETYL-BETA-D-GLUCOSAMINE.  
CC -!- PATHWAY: LIPOPOLYSACCHARIDE BIOSYNTHESIS.  
CC -!- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 52.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M94855; AA24979.1; -;  
DR EMBL; U32842; AAC23345.1; -;  
DR TIGR; H11699; -;  
KW Transferase; Glycosyltransferase; Lipopolysaccharide biosynthesis;  
KW Complete proteome.  
FT CONFLICT 4 4 M -> I (IN REF. 1).  
FT CONFLICT 68 68 F -> S (IN REF. 1).  
FT CONFLICT 130 130 G -> D (IN REF. 1).  
FT CONFLICT 220 220 Y -> C (IN REF. 1).  
SQ SEQUENCE 304 AA; 35706 MW; B6D03890AC1CDD28 CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 304;  
Best Local Similarity 71.4%; Pred. No. 1.4e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 5 QOFFGLM 11  
DB 25 EQFFGV 31  
:||||:|  
  
RESULT 98  
RADA\_METVO STANDARD; PRT; 322 AA.  
ID RADA\_METVO  
AC O73948;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DNA REPAIR AND RECOMBINATION PROTEIN RADA.  
GN RADA.  
OS Methanococcus voltae.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
OX NCBI\_TaxID=2188;  
[1]  
SEQUENCE FROM N.A.  
RA STRAIN-PS / DSM 1537;  
RA Reich C.I., Buldak G.L., McNeill L.K.;  
RT "The Rada protein from the archaeon *Methanococcus voltae*: a functional  
RT homolog of eukaryotic Rad51.";  
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
[2]  
SEQUENCE OF 112-254 FROM N.A.  
RX MEDLINE=99121030; PubMed=9922255;  
RA Sandler S.J., Hugenholtz P., Schleper C., DeLong E.F., Pace N.R.,  
RA Clark A.J.;  
RT "Diversity of rada genes from cultured and uncultured archaea:  
RT comparative analysis of putative Rada proteins and their use as a  
RT phylogenetic marker.";  
RL J. Bacteriol. 181:907-915(1999).  
CC -!- FUNCTION: INVOLVED IN DNA REPAIR AND IN HOMOLOGOUS RECOMBINATION.  
CC BINDS AND ASSEMBLES ON SINGLE-STRANDED DNA TO FORM A NUCLEOPROTEIN  
CC FILAMENT. HYDROLYZES ATP IN A SSNA-DEPENDENT MANNER AND PROMOTES  
CC DNA STRAND EXCHANGE BETWEEN HOMOLOGOUS DNA MOLECULES (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE EUKARYOTIC RECA-LIKE PROTEIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

```
CC -----
DR EMBL; AF008421; AAC23499.1; -.
DR EMBL; AF090200; AAD16066.1; -.
DR InterPro; IPR000445; HHH.
DR InterPro; IPR001553; RecA.
DR InterPro; IPR003583; HHH_1.
DR SMART; SM00278; HHH1; 2.
DR PROSITE; PS0162; RECA_2; 1.
DR PROSITE; PS0163; RECA_3; 1.
KW DNA damage; DNA recombination; ATP-binding; DNA-binding.
FT NP_BIND 105 112 ATP (POTENTIAL).
SQ SEQUENCE 322 AA; 35189 MW; 8A9F5EFB927344B6 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 322;
Best Local Similarity 62.5%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 PKQOFFGL 10
Db 261 KPDAFFGM 268

RESULT 99
VPRT_SMRVH STANDARD; PRT; 323 AA.
AC P21407;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE PROTEASE (EC 3.4.23.-).
GN PRT.
OS Squirrel monkey retrovirus (SMRV-H) (SMRV-HLB).
OC Viruses; Retroviral viruses; Retroviridae; Betaretrovirus.
OX NCBI_TaxID=11856;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89073750; PubMed=3201749;
RA Oda T., Ikeda S., Watanabe S., Hatsushika M., Akiyama K.,
RA Mitsunobu F.;
RT "Molecular cloning, complete nucleotide sequence, and gene structure
RT of the provirus genome of a retrovirus produced in a human
RT lymphoblastoid cell line.";
RL Virology 167:468-476(1988).
CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE
CC RETROPEPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M23385; AAA66452.1; ALT_INIT.
CC PIR; B31827; PRLJHD.
CC HSP; P03366; 9HVP.
DR InterPro; IPR001995; Asp_prot_retrov.
DR InterPro; IPR001969; Asp_protease.
DR InterPro; IPR000467; G_patch.
DR InterPro; IPR001428; dufPase.
DR Pfam; PF00692; dufPase; 1.
DR Pfam; PF01585; G_patch; 1.
DR Pfam; PF00077; rvp; 1.
DR ProDom; PD000946; dufPase; 1.
DR SMART; SM00443; G_patch; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
DR PROSITE; PS00175; ASP_PROT_RETROV; 1.
KW Hydrolase; Aspartyl protease.
FT ACT_SITE 193 193 BY SIMILARITY.
SQ SEQUENCE 323 AA; 35126 MW; 5D6CEA38BA932786 CRC64;
```

```
Query Match 50.8%; Score 31; DB 1; Length 323;
Best Local Similarity 60.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKQOFFGLM 11
Db 82 PLPPQTEGLI 91

RESULT 100
PUR5_METH STANDARD; PRT; 338 AA.
ID PUR5_METH
AC O27272;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE PHOSPHORIBOSYLGLYCINAMIDINE CYCLO-LIGASE (EC 6.3.3.1)
DE (AIRS) (PHOSPHORIBOSYL-AMINOIMIDAZOLE SYNTHETASE) (AIR SYNTHASE).
GN PURM OR MTH1204.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC Methanothermobacter.
OX NCBI_TaxID=145262;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DELTA H.
RX MEDLINE=98037514; PubMed=9371463;
RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwanji N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Nolling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -! CATALYTIC ACTIVITY: ATP + 5'-PHOSPHORIBOSYL-5-AMINOIMIDAZOLE
CC ADP + ORTHOPHOSPHATE + 5'-PHOSPHORIBOSYL-FORMYLGLYCINAMIDINE -
CC (BY SIMILARITY).
CC -! PATHWAY: FIFTH STEP IN DE NOVO PURINE BIOSYNTHESIS
CC (BY SIMILARITY).
CC -! SIMILARITY: TO OTHER AIRS FROM BACTERIA AND EUKARYOTES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000888; AAB85693.1; -.
CC InterPro; IPR000728; AIRS_related.
CC Pfam; PF00586; AIRS; 1.
KW Purine biosynthesis; Ligase; Complete proteome.
SQ SEQUENCE 338 AA; 35983 MW; 5C6AA2B6562E0E17 CRC64;
```

```
Query Match 50.8%; Score 31; DB 1; Length 338;
Best Local Similarity 71.4%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQOFF 8
Db 264 PEPQQIF 270

RESULT 101
PYRD_PASMU STANDARD; PRT; 339 AA.
ID PYRD_PASMU
AC P57858;
```

20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE DIHYDROOROTATE DEHYDROGENASE (EC 1.3.3.1) (DIHYDROOROTATE OXIDASE)  
DE (DHODPHASE) (DHODASE) (DHOD).  
GN PYRD OR PM0617.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PM70;  
RX MEDLINE=21145866; PubMed=11248100;  
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -1- CATALYTIC ACTIVITY: L-DIHYDROOROTATE + O(2) -> OROTATE + H(2)O(2).  
CC -1- COFACTOR: FMN (BY SIMILARITY).  
CC -1- PATHWAY: FOURTH STEP IN PYRIMIDINE BIOSYNTHESIS.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: INNER SIDE OF THE MEMBRANE (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE DIHYDROOROTATE DEHYDROGENASE FAMILY.  
CC SUBFAMILY 2.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; AE006098; AAK02701.1; -.  
DR InterPro; IPR001295; DHO\_dh.  
DR InterPro; IPR003009; FMN\_enzyme.  
DR Pfam; PF01180; DHODehase; 1.  
DR PROSITE; PS00911; DHODPHASE\_1; 1.  
DR PROSITE; PS00912; DHODPHASE\_2; 1.  
KW Pyrimidine biosynthesis; Oxidoreductase; Flavoprotein; FMN;  
FT NP\_BIND 292 300 FMN (POTENTIAL).  
FT COMPLETE PROTEOME.  
SQ SEQUENCE 339 AA; 36919 MW; AEA5E07B29942D68 CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 339;  
Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 PKPOOF 7  
| | | | |  
DB 98 PKPROF 103  
  
RESULT 102  
CPXE\_STRGO STANDARD; PRT; 405 AA.  
AC P18326;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE CYTOCHROME P450-SUL (EC 1.14.-.-) (P450-CVAL) (CYP105A1).  
GN CYP105A1 OR SUAC.  
OS Streptomyces griseolus.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1909;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-31.  
RC STRAIN=ATCC 11796;  
RX MEDLINE=90264332; PubMed=2345149;  
RA Omer C.A., Lenstra R., Little P.J., Dean C., Tepperman J.M.,  
RA Leto K.J., Romesser J.A., O'Keefe D.P.;

"Genes for two herbicide-inducible cytochromes P-450 from  
Streptomyces griseolus.";  
J. Bacteriol. 172:3335-3345(1990).  
CC -1- FUNCTION: METABOLISM OF A NUMBER OF SULFONYLUREA HERBICIDES.  
CC -1- INDUCTION: BY HERBICIDE.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; M32238; AAA26823.1; -.  
DR PIR; A35401; A35401.  
DR HSSP; P23295; 2ROM.  
DR InterPro; IPR001128; Cyt\_P450.  
DR Pfam; PF00067; P450; 1.  
DR PRINTS; PR00359; BP450.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
KW Oxidoreductase; Monooxygenase; Electron transport; Heme.  
FT INIT\_MET 0  
FT BINDING 354 354 HEME (BY SIMILARITY).  
SQ SEQUENCE 405 AA; 44081 MW; 92AB36E064FD0B3E CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 405;  
Best Local Similarity 60.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 RPKQPFGL 10  
| | | | |  
DB 88 RESPQAFGL 97  
  
RESULT 103  
TIG\_MYCTU STANDARD; PRT; 466 AA.  
ID TIG\_MYCTU  
AC O53189;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE TRIGGER FACTOR (TF).  
GN TIG OR RV2462C OR MT2537 OR MTV008.18C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Horsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains.";  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: INVOLVED IN PROTEIN EXPORT. ACTS AS A CHAPERONE BY MAINTAINING THE NEWLY SYNTHESIZED PROTEIN IN AN OPEN CONFORMATION (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE FKBP-TYPE PPIASE FAMILY. TIG SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AL021246; CAA16039.1; -  
DR EMBL: AB007090; AAK46837.1; -  
DR TIGR: MT2537; -  
DR TubercuList; RV2462c; -  
DR InterPro: IPR001179; FKBP\_PPIase.  
DR PROSITE: PS00453; FKBP\_PPIase\_1; FALSE\_NEG.  
DR PROSITE: PS00454; FKBP\_PPIase\_2; FALSE\_NEG.  
DR PROSITE: PS00059; FKBP\_PPIase\_3; 1.  
KW Cell division; Chaperone; Isomerase; Rotamase; Complete proteome.  
FT DOMAIN 162 243 PPIASE, FKBP-TYPE.  
SQ SEQUENCE 466 AA; 50616 MW; AFF5DC88976036D CRC64;

Query Match 50.8%; Score 31; DB 1; Length 466;  
Best Local Similarity 55.8%; Pred. NO. 2.2e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 3 KPQOFFGLM 11  
:|:|:|:|:  
Db 387 EPQOLFGL 395

RESULT 104  
TRE2\_SYNY3  
ID TRE2\_SYNY3 STANDARD; PRT; 485 AA.  
AC P74130;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ANTHRANILATE SYNTHASE COMPONENT I-LIKE PROTEIN (EC 4.1.3.27).  
GN TRPE2 OR SLR1979.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y., Miyajima N., Hiroseawa M., Sugliura M., Sasamoto S., Kimura T., Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S., Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
CC -1- CATALYTIC ACTIVITY: CHORISMATE + L-GLUTAMINE = ANTHRANILATE + PYRUVATE + L-GLUTAMATE.  
CC -1- PATHWAY: FIRST STEP IN BIOSYNTHESIS OF TRYPTOPHAN.  
CC -1- SUBUNIT: Tetramer of two components I and two components II (by similarity).  
CC -1- MISCELLANEOUS: COMPONENT I CATALYZES THE FORMATION OF ANTHRANILATE USING AMMONIA RATHER THAN GLUTAMINE, WHEREAS COMPONENT II PROVIDES GLUTAMINE AMIDOTRANSFERASE ACTIVITY.  
CC -1- SIMILARITY: BELONGS TO THE ANTHRANILATE SYNTHASE COMPONENT I FAMILY.  
CC -1- CAUTION: THIS IS A DIVERGENT FORM OF TRPE. IT IS NOT OBVIOUS IF IT

CC IS ACTIVE IN TRP BIOSYNTHESIS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D90912; BAA18216.1; -  
DR InterPro: IPR000350; Chorismate\_bind.  
DR Pfam: PF00425; Chorismate\_bind; 1.  
DR PRINTS: PR00095; ANTSNTHASE1.  
DR ProDom: PD000779; Chorismate\_bind; 1.  
KW Tryptophan biosynthesis; Lyase; Complete proteome.  
SQ SEQUENCE 485 AA; 54270 MW; 4F25ECCB3857BC7C CRC64;  
Query Match 50.8%; Score 31; DB 1; Length 485;  
Best Local Similarity 40.0%; Pred. NO. 2.2e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
QY 2 KPQOFFGLM 11  
:|:|:|:|:  
Db 98 KPPEEIFSFL 107  
RESULT 105  
EXON\_HSV6U  
ID EXON\_HSV6U STANDARD; PRT; 488 AA.  
AC P24447;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ALKALINE EXONUCLEASE (EC 3.1.11.-).  
GN U70 OR 16R.  
OS Human herpesvirus (type 6 / strain Uganda-1102) (HHV6).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Roseolovirus.  
OX NCBI\_TaxID=10370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90080132; PubMed=2152817;  
RA Lawrence G.L., Chee M., Craxton M.A., Gompels U.A., Honess R.W., Barrell B.G.;  
RT "Human herpesvirus 6 is closely related to human cytomegalovirus.";  
RL J. Virol. 64:287-299(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95266321; PubMed=7747482;  
RA Gompels U.A., Nicholas J., Lawrence G., Jones M., Thomscn B.J., Martin M.E., Efsthliou S., Craxton M., Macaulay H.A.;  
RT "The DNA sequence of human herpesvirus-6: structure, coding content, and genome evolution.";  
RL Virology 209:29-51(1995).  
CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES ALKALINE EXONUCLEASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X83413; CAA58362.1; -  
DR EMBL: M68963; AAA65578.1; -  
DR PIR: F36769; QOBEHS.  
DR InterPro: IPR001616; Herpes\_alk\_exo.  
DR Pfam: PF01771; Herpes\_alk\_exo; 1.  
DR PRINTS: PR00924; ALKEXNUCLASE.



KW Hydrolase; Nuclease; Exonuclease.  
SQ SEQUENCE 488 AA; 56644 MW; 0F38A10597366A5B CRC64;

Query Match 50.8%; Score 31; DB 1; Length 488;  
Best Local Similarity 85.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFFG 9  
| | | | |  
DB 161 KGQOQFFG 167

## RESULT 106

ID EXON\_HSV6Z STANDARD; PRT; 488 AA.  
AC P52448;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ALKALINE EXONUCLEASE (EC 3.1.11.-).  
GN U70 OR CH3R.  
OS Human herpesvirus (type 6 / strain 229) (HHV6).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Roseolovirus.  
OX NCBI\_TaxID=36351;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=96195263; PubMed=8634027;  
RA Lindquester G.J., Inoue N., Allen R.D., Castelli J.W.,  
RA Stamey F.R., Dambaugh T.R., O'Brian J.J., Danovich R.M.,  
RA Frenkel N., Pellett P.E.;  
RT "Restriction endonuclease mapping and molecular cloning of the human  
herpesvirus 6 variant B strain 229 genome.";  
RL Arch. Virol. 141:367-379(1996).  
CC -!- SIMILARITY: BELONGS TO THE HERPESVIRUSES ALKALINE EXONUCLEASE  
FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AF157706; AAB06353.1; -  
CC InterPro; IPR001616; Herpes\_alk\_exo.  
CC Pfam; PF01771; Herpes\_alk\_exo; 1.  
CC PRINTS; PR00924; ALKEXNUCLASE.  
KW Hydrolase; Nuclease; Exonuclease.  
SQ SEQUENCE 488 AA; 56687 MW; AE2872028D4B3D90 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 488;  
Best Local Similarity 85.7%; Pred. No. 2.3e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFFG 9  
| | | | |  
DB 161 KGQOQFFG 167

## RESULT 107

ID SYK\_MYCHO STANDARD; PRT; 488 AA.  
AC P46191;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE LYSYL-TRNA SYNTHETASE (EC 6.1.1.6) (LYSINE--TRNA LIGASE) (LYSRS).  
GN LYS5.  
OS Mycoplasma hominis.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2098;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PG21;  
RA MEDLINE=94237425; PubMed=8181699;  
RA Ozkokmen D., Birkelund S., Christiansen G.;  
RT "Characterization of a Mycoplasma hominis gene encoding lysyl-trna  
synthetase (LysRS).";  
RL FEMS Microbiol. Lett. 116:277-282(1994).  
CC -!- CATALYTIC ACTIVITY: ATP + L-LYSINE + TRNA(LYS) = AMP +  
PYROPHOSPHATE + L-LYSYL-TRNA(LYS).  
CC -!- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -!- SIMILARITY: BELONGS TO CLASS-II AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X74912; CAA52877.1; -  
CC HSP; P14825; LYL.  
CC InterPro; IPR002106; AA\_trna\_ligase\_II.  
CC InterPro; IPR002309; trna-synt\_2.  
CC InterPro; IPR002313; trna-synt\_lys\_2.  
CC Pfam; PF00152; trna-synt\_2; 1.  
CC Pfam; PF01336; trna-anti\_1.  
CC PRINTS; PR00982; TRNASYNTHLYS.  
CC PROSITE; PS00179; AA\_TRNA\_LIGASE\_II\_1; 1.  
CC PROSITE; PS00339; AA\_TRNA\_LIGASE\_II\_2; FALSE\_NEG.  
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding.  
SQ SEQUENCE 488 AA; 56868 MW; 83D34BF37E21E32E CRC64;

Query Match 50.8%; Score 31; DB 1; Length 488;  
Best Local Similarity 45.5%; Pred. No. 2.3e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RKPKQOQFFGLM 11  
| | | | |  
DB 137 KPLPDKFHGLV 147

## RESULT 108

ID CPVL\_BRARE STANDARD; PRT; 509 AA.  
AC O42145;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE CYTOCHROME P450 19 (AROMATASE) (EC 1.14.14.1) (CYP19) (ESTROGEN  
SYNTHETASE) (P-450AROM).  
GN CYP19.  
OS Brachydanio rerio (zebrafish) (zebra danio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
OC Cypriniformes; Cyprinidae; Rasbora; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Bauer M.P., Goetz F.W.;  
RT "Isolation and characterization of a zebrafish (Danio rerio) aromatase  
cDNA.";  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: CATALYZES THE FORMATION OF AROMATIC C18 ESTROGENS FROM  
C19 ANDROGENS (BY SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: RH + REDUCED FLAVOPROTEIN + O(2) = ROH +

CC OXIDIZED FLAVOPROTEIN + H(2)O.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND.  
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF004521; AAB65788.1; -  
CC ZFIN: ZDB-GENE-990415-43; cyp19.  
CC InterPro: IPR001128; Cyt\_P450.  
CC Pfam: PF00067; P450; 1.  
CC PRINTS: PR00385; P450.  
CC PROSITE: PS00086; CYTOCHROME\_P450; 1.  
KW Electron transport; Oxidoreductase; Monooxygenase; Membrane;  
KW Heme.  
FT BINDING 456 456 HEME (BY SIMILARITY).  
FT SEQUENCE 509 AA; 57354 MW; F04532DEA8FDB628 CRC64;  
SQ  
  
Query Match 50.8%; Score 31; DB 1; Length 509;  
Best Local Similarity 66.7%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PKPQQFGL 10  
| | | | |  
DB 68. PGPSFFGL 76  
  
RESULT 109  
DHAF\_VIBHA STANDARD; PRT; 510 AA.  
AC Q56694;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE FATTY ALDEHYDE DEHYDROGENASE (EC 1.2.1.3).  
GN ALDH.  
OS Vibrio harveyi.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=669;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC STRAIN=B392;  
RX MEDLINE=96118391; PubMed=8527447;  
RA Vedadi M., Zittner R., Smillie L., Meighen E.;  
RT "Involvement of cysteine 289 in the catalytic activity of an NADP(+) -  
RT specific fatty aldehyde dehydrogenase from Vibrio harveyi.";  
RL Biochemistry 34:16725-16732(1995).  
RN [2]  
RN X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).  
RC STRAIN=B392;  
RX MEDLINE=20363527; PubMed=10903148;  
RA Ahvazi B., Coulombe R., Delarge M., Vedadi M., Zhang L., Meighen E.,  
RA Vrielink A.;  
RT "Crystal structure of the NADP+-dependent aldehyde dehydrogenase from  
RT Vibrio harveyi: structural implications for cofactor specificity and  
RT affinity.";  
RL Blochem. J. 349:853-861(2000).  
CC -1- FUNCTION: CATALYZES THE OXIDATION OF LONG-CHAIN ALIPHATIC  
CC ALDEHYDES TO ACIDS. MAY BE IMPLICATED IN CONTROLLING LUMINESCENCE  
CC AS IT CATALYZES THE OXIDATION OF THE FATTY ALDEHYDE SUBSTRATE FOR  
CC THE LIGHT-EMITTING REACTION.  
CC -1- CATALYTIC ACTIVITY: ALDEHYDE + NADP(+) + H(2)O = ACID + NADPH.  
CC -1- COFACTOR: HAS A HIGH SPECIFICITY FOR NADP.  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SIMILARITY: BELONGS TO THE ALDEHYDE DEHYDROGENASES FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: U39638; AAA89078.1; -  
CC PDB: 1EYV; 24-MAY-00.  
CC PDB: 1E20; 24-MAY-00.  
CC InterPro: IPR002086; Aldehyde\_dehydr.  
CC Pfam: PF00171; aldehyd; 1.  
CC PROSITE: PS00070; ALDEHYDE-DEHYDR\_CYS; FALSE\_NEG.  
CC PROSITE: PS00687; ALDEHYDE-DEHYDR\_GLU; FALSE\_NEG.  
KW Oxidoreductase; NADP; 3D-structure.  
FT NE\_BIND 229 234 NADP (ADP PART).  
FT ACT\_SITE 253 253  
FT ACT\_SITE 289 289  
SQ SEQUENCE 510 AA; 54459 MW; E132F2406AA3F47A CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 510;  
Best Local Similarity 55.6%; Pred. No. 2.4e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 RPKPQQFFG 9  
| | | | |  
DB 244 RPEPIPYG 252  
  
RESULT 110  
YTE4\_CAEEL STANDARD; PRT; 546 AA.  
AC Q17865;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHETICAL 60.9 KDA PROTEIN C09G1.4 IN CHROMOSOME X.  
GN C09G1.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA McMurray A.;  
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP REVISIONS.  
RA Jones S.J.M.;  
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: Z50176; CAA90541.1; -  
CC WormPep: C09G1.4; CE08042.  
CC InterPro: IPR001478; PDZ.  
CC SMART: SM00228; PDZ; 1.  
CC PROSITE: PS50106; PDZ; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 546 AA; 60864 MW; 2764EBB62461E3C2 CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 546;  
Best Local Similarity 63.6%; Pred. No. 2.5e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
||| |||  
Db 314 RPLPQQQTSLM 324

RESULT 111  
GPCL1\_RAT  
ID GPCL1\_RAT STANDARD; PRT; 558 AA.  
AC P35053;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE GLYPICAN-1 PRECURSOR (HSPG M12).  
GN GPCL1.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 24-55 AND 424-445.  
RC TISSUE=Brain.  
RX MEDLINE=93038690; PubMed=14117860;  
RA Karthikeyan L., Maurel P., Rauch U., Margolis R.K., Margolis R.U.;  
RT "Cloning of a major heparan sulfate proteoglycan from brain and  
RT identification as the rat form of glypican".  
RL Biochem. Biophys. Res. Commun. 188:395-401(1992).  
RN [2]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 83-112; 196-207 AND 422-443.  
RC STRAIN=NEW ENGLAND DECONESS HOSPITAL;  
RX MEDLINE=94267529; PubMed=8207484;  
RA Litwack E.D., Stipp C.S., Kumbasar A., Lander A.D.;  
RT "Neuronal expression of glypican, a cell-surface  
RT glycosylphosphatidylinositol-anchored heparan sulfate proteoglycan,  
RT in the adult rat nervous system".  
RL J. Neurosci. 14:3713-3724(1994).  
CC -1- FUNCTION: CELL SURFACE PROTEOGLYCAN THAT BEARS HEPARAN SULFATE.  
CC MAY PLAY AN IMPORTANT ROLE IN THE TROPIC AND INJURY RESPONSES OF  
CC NEURONS.  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A GPI-ANCHOR.  
CC -1- TISSUE SPECIFICITY: NERVOUS SYSTEM.  
CC -1- PFM: THIS CELL-ASSOCIATED GLYPICAN IS FURTHER PROCESSED TO GIVE  
CC RISE TO A MEDIUM-RELEASED SPECIES.  
CC -1- SIMILARITY: BELONGS TO THE GLYPICAN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L02896; AAA86439.1; -  
DR EMBL; L34067; AAA41251.1; -  
DR PIR; JC1281; JC1281.  
DR InterPro: IPR001863; Glypican.  
DR Pfam; PF01153; Glypican; 1.  
DR PROSITE; PS01207; GLYPICAN; 1.  
KW Proteoglycan; Heparan sulfate; Glycoprotein; Signal; GPI-anchor;  
KW Extracellular matrix.  
FT SIGNAL 1 23  
FT CHAIN 24 530 GLYPICAN-1..  
FT PROPEP 531 558 REMOVED IN MATURE FORM (POTENTIAL).  
FT LIPID 530 530 GPI-ANCHOR (POTENTIAL).  
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 116 116 O-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 55 55 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).  
FT CARBOHYD 486 486 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).  
FT CARBOHYD 488 488 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).  
FT CARBOHYD 490 490 O-LINKED (GLYCOSAMINOGLYCAN) (POTENTIAL).  
FT CONFLICT 21 21 T -> A (IN REF. 2).  
FT CONFLICT 312 312 Y -> N (IN REF. 2).

FT CONFLICT 362 362 A -> G (IN REF. 2).  
FT CONFLICT 437 437 I -> G (IN REF. 2; AA SEQUENCE).  
FT CONFLICT 443 443 E -> D (IN REF. 2; AA SEQUENCE).  
FT CONFLICT 515 515 I -> T (IN REF. 2).  
SQ SEQUENCE 558 AA; 61734 MW; E2878A854B9A1D7F CRC64;

Query Match 50.8%; Score 31; DB 1; Length 558;  
Best Local Similarity 62.5%; Pred.No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQFF 8  
||:| ||  
Db 534 REPHYFF 541

RESULT 112  
MM02\_HUMAN  
ID MM02\_HUMAN STANDARD; PRT; 660 AA.  
AC P08253;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE 72 KDA TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KDA  
DE GELATINASE) (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A)  
DE (TBE-1).  
DE MMP2 OR CLG4A.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE OF 19-660 FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=86198218; PubMed=2834383;  
RA Collier I.E., Wilhelm S.M., Eisen A.Z., Marmer B.L., Grant G.A.,  
RA Seltzer J.L., Kronberger A., He C., Bauer E.A., Goldberg G.I.;  
RT "H-ras oncogene-transformed human bronchial epithelial cells (TBE-1)  
RT secrete a single metalloprotease capable of degrading basement  
RT membrane collagen".  
RL J. Biol. Chem. 263:6579-6587(1988).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91236162; PubMed=1851724;  
RA Collier I.E., Bruns G.A.P., Goldberg G.I., Gerhard D.S.;  
RT "On the structure and chromosome location of the 72- and 92-kDa human  
RT type IV collagenase genes".  
RL Genomics 9:429-434(1991).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90293047; PubMed=2162831;  
RA Huhtala P., Chow L.T., Tryggvason K.;  
RT "Structure of the human type IV collagenase gene".  
RL J. Biol. Chem. 265:11077-11082(1990).  
RN [4]  
RP SEQUENCE OF 1-51 FROM N.A.  
RX MEDLINE=90228972; PubMed=2158484;  
RA Huhtala P., Eddy R.L., Fan Y.S., Byers M.G., Shows T.B.,  
RA Tryggvason K.;  
RT "Completion of the primary structure of the human type IV collagenase  
RT preproenzyme and assignment of the gene (CLG4) to the q21 region of  
RT chromosome 16".  
RL Genomics 6:554-559(1990).  
RN [5]  
RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 443-660.  
RX MEDLINE=96069777; PubMed=7583664;  
RA Libson A.M., Gittis A.G., Collier I.E., Marmer B.L., Goldberg G.I.,  
RA Lattman E.E.;  
RT "Crystal structure of the haemopexin-like C-terminal domain of  
RT gelatinase A".  
RL Nat. Struct. Biol. 2:938-942(1995).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 458-660.  
RX MEDLINE=96140723; PubMed=8549817;

RA Gohlke U., Gomis-Ruth F.X., Crabbe T., Murphy G., Docherty A.J.,  
RA Bode W.;  
RT "The C-terminal (haemopexin-like) domain structure of human  
RT gelatinase A (MMP2): structural implications for its function.";  
RL FEBS Lett. 378:126-130(1996).  
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES  
CC IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLY-|-  
CC ILE-ALA-GLY-GLN.  
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.  
CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.  
CC -1- TISSUE SPECIFICITY: PRODUCED BY NORMAL SKIN FIBROBLASTS.  
CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; J03210; AAA35701.1; -  
DR EMBL; M33789; AAA52027.1; -  
DR EMBL; M55593; AAA52028.1; -  
DR EMBL; M58552; AAA52028.1; JOINED.  
DR EMBL; M55582; AAA52028.1; JOINED.  
DR EMBL; M55583; AAA52028.1; JOINED.  
DR EMBL; M55584; AAA52028.1; JOINED.  
DR EMBL; M55585; AAA52028.1; JOINED.  
DR EMBL; M55586; AAA52028.1; JOINED.  
DR EMBL; M55587; AAA52028.1; JOINED.  
DR EMBL; M55588; AAA52028.1; JOINED.  
DR EMBL; M55589; AAA52028.1; JOINED.  
DR EMBL; M55590; AAA52028.1; JOINED.  
DR EMBL; M55591; AAA52028.1; JOINED.  
DR EMBL; M55592; AAA52028.1; JOINED.  
DR PIR; A28153; A28153.  
DR PDB; 1RG6; 10-JUN-96.  
DR PDB; 1GEN; 17-AUG-96.  
DR MEROPS; M10.003; -  
DR MTM; 120360; -  
DR InterPro; IPR000562; FN\_Type\_II.  
DR InterPro; IPR000585; Hemopexin.  
DR InterPro; IPR001818; Matrixin.  
DR InterPro; IPR000130; Zn\_MTPeptidse.  
DR Pfam; PF00040; fn2; 3.  
DR Pfam; PF00045; hemopexin; 4.  
DR Pfam; PF00413; Peptidase\_M10; 1.  
DR PRINTS; PR00013; FNTYPEII.  
DR PRINTS; PR00138; MATRIXIN.  
DR ProDom; PD000995; FN\_Type\_II; 3.  
DR SMART; SM00059; FN2; 3.  
DR SMART; SM00120; HX; 4.  
DR SMART; SM00235; ZnMc; 1.  
DR PROSITE; PS00023; FIBRONECTIN\_2; 3.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
DR PROSITE; PS00346; CYSTEINE\_SWITCH; 1.  
KW Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;  
KW Collagen degradation; Extracellular matrix; Repeat; Signal;  
KW 3D-structure.  
FT SIGNAL 1 29 POTENTIAL.  
FT PROPEP 30 109 ACTIVATION PEPTIDE.  
FT CHAIN 110 660 72 KDA TYPE IV COLLAGENASE.  
FT DOMAIN 110 221 COLLAGENASE-LIKE.  
FT DOMAIN 222 396 COLLAGEN-BINDING.  
FT DOMAIN 397 465 COLLAGENASE-LIKE.  
FT DOMAIN 226 283 FIBRONECTIN TYPE-II 1.  
FT DOMAIN 284 341 FIBRONECTIN TYPE-II 2.  
FT DOMAIN 342 399 FIBRONECTIN TYPE-II 3.

FT DOMAIN 466 660 HEMOPEXIN-LIKE.  
FT SITE 102 102 CYSTEINE SWITCH (POTENTIAL).  
FT METAL 403 403 ZINC (CATALYTIC) (BY SIMILARITY).  
FT ACT\_SITE 404 404 BY SIMILARITY.  
FT METAL 407 407 ZINC (CATALYTIC) (BY SIMILARITY).  
FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).  
FT CARBOHYD 573 573 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 642 642 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 469 660  
SQ SEQUENCE 660 AA; 73882 MW; BC7147DC8B49F289 CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 660;  
Best Local Similarity 75.0%; Pred. No. 3.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 KPOQFFGL 10  
| :|||  
Db 76 KMQKFFGL 83  
  
RESULT 113  
MM02\_MOUSE STANDARD; PRT; 662 AA.  
ID MM02\_MOUSE  
AC P33434;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE 72 KDA TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KDA  
DE GELATINASE) (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A).  
GN MMP2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=92218452; PubMed=1373140;  
RA Reponen P., Sahlberg C., Huhtala P., Hurskainen T., Thesleff I.,  
RA Tryggvason K.;  
RT "Molecular cloning of murine 72-kDa type IV collagenase and its  
RT expression during mouse development.";  
RL J. Biol. Chem. 267:7856-7862(1992).  
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES  
CC IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLY-|-  
CC ILE-ALA-GLY-GLN.  
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.  
CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.  
CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M84324; AAA39338.1; -  
DR PIR; A42496; A42496.  
DR HSSP; P08253; 1RTG.  
DR MEROPS; M10.003; -  
DR MGD; MGI:97009; Mmp2.  
DR InterPro; IPR000562; FN\_Type\_II.  
DR InterPro; IPR000585; Hemopexin.  
DR InterPro; IPR001818; Matrixin.  
DR InterPro; IPR000130; Zn\_MTPeptidse.  
DR Pfam; PF00040; fn2; 3.  
DR Pfam; PF00045; hemopexin; 4.  
DR Pfam; PF00413; Peptidase\_M10; 1.

```
DR PRINTS; PR000138; FNTypeII.
DR PRINTS; PR000138; FNTypeII.
DR PRODOM; PR000095; FNTypeII; 3.
DR SMART; SM000059; FN2; 3.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZNMC; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 3.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
KW Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
collagen degradation; Extracellular matrix; Repeat; Signal.
FT SIGNAL 1 29
FT PROPEP 30 109 ACTIVATION PEPTIDE.
FT CHAIN 110 662 72 KDA TYPE IV COLLAGENASE.
FT DOMAIN 110 221 COLLAGENASE-LIKE.
FT DOMAIN 222 396 COLLAGENASE-LIKE.
FT DOMAIN 397 467 COLLAGENASE-LIKE.
FT DOMAIN 226 283 FIBRONECTIN TYPE-II 1.
FT DOMAIN 284 341 FIBRONECTIN TYPE-II 2.
FT DOMAIN 342 399 FIBRONECTIN TYPE-II 3.
FT DOMAIN 468 662 HEMOPEXIN-LIKE.
FT SITE 102 102 CYSTEINE SWITCH (POTENTIAL).
FT METAL 403 404 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 404 404 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 407 407 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 575 575 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 644 644 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 471 662 BY SIMILARITY.
SQ SEQUENCE 662 AA; 74102 MW; C630A7DBDB272F02 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 662;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFFGL 10
DB 76 KMQKFFGL 83

RESULT 114
MM02_RABIT STANDARD; PRT; 662 AA.
AC P50757;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 72 KDA TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KDA
DE GELATINASE) (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A).
GN MMP2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAPANESE WHITE; TISSUE=Articular joint;
RA MEDLINE=96283805; PubMed=8679695;
RA Matsumoto S., Katoh M., Watanabe T., Masuho Y.;
RT "Molecular cloning of rabbit matrix metalloproteinase-2 and its broad
RT expression at several tissues."
RL Biochim. Biophys. Acta 1307:137-139(1996).
CC -! CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES
CC IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLY-|-
CC ILE-ALA-GLY-GLN.
CC -! COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.
CC -! SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.
CC -! SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.
CC -! SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.
CC -! SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC
CC METALLOPROTEASE) ALSO KNOWN AS MATRIXIN SUBFAMILY.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D63579; BAA09796.1; -.
DR HSSP; P08253; IRTG.
DR MEROPS; M10.003; -.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR000585; Hemoexin.
DR InterPro; IPR001818; Matrixin.
DR InterPro; IPR000130; Zn_Mtpeptdse.
DR Pfam; PF00040; fn2; 3.
DR Pfam; PF00045; hemoexin; 4.
DR Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PR00013; FNYPEII.
DR PRINTS; PR00138; MATRINX.
DR ProDom; PR000995; FN_Type_II; 3.
DR SMART; SM00059; FN2; 3.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZNMC; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 3.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
KW Hydrolase; Metalloprotease; Glycoprotein; Zinc; Zymogen; Calcium;
collagen degradation; Extracellular matrix; Repeat; Signal.
FT SIGNAL 1 29
FT PROPEP 30 109 ACTIVATION PEPTIDE.
FT CHAIN 110 662 72 KDA TYPE IV COLLAGENASE.
FT DOMAIN 110 221 COLLAGENASE-LIKE.
FT DOMAIN 222 396 COLLAGENASE-LIKE.
FT DOMAIN 397 467 COLLAGENASE-LIKE.
FT DOMAIN 226 283 FIBRONECTIN TYPE-II 1.
FT DOMAIN 284 341 FIBRONECTIN TYPE-II 2.
FT DOMAIN 342 399 FIBRONECTIN TYPE-II 3.
FT DOMAIN 468 662 HEMOPEXIN-LIKE.
FT SITE 102 102 CYSTEINE SWITCH (POTENTIAL).
FT METAL 403 404 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 404 404 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 407 407 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 413 413 ZINC (CATALYTIC) (BY SIMILARITY).
FT CARBOHYD 575 575 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 644 644 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 471 662 BY SIMILARITY.
SQ SEQUENCE 662 AA; 73803 MW; 1CC246B270E440C8 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 662;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFFGL 10
DB 76 KMQKFFGL 83

RESULT 115
MM02_RAT STANDARD; PRT; 662 AA.
AC P33436; P97581;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 72 KDA TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KDA
DE GELATINASE) (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A).
GN MMP2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93249363; PubMed=7916617;  
RA Marti H.P., McNeil L., Davies M., Martin J., Lovett D.H.;  
RT "Homology cloning of rat 72 kDa type IV collagenase: cytokine and  
second-messenger inducibility in glomerular mesangial cells.";  
RL Biochem. J. 291:441-446(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WISTAR; TISSUE=Skin;  
RA Okada A., Basset P.;  
RT "The cloning of the cDNA encoding rat gelatinase A from a rat skin  
tumor cDNA library";  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES  
IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLY-|-  
ILE-ALA-GLY-GLN.  
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.  
CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3  
CC -1- SIMILARITY: CONTAINS 1 HEMOPLEXIN-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X71466; CNA50583.1; -  
DR EMBL; U65656; AAB41692.1; -  
DR PIR; S34780; S34780.  
DR HSSP; P08253; IRTG.  
DR MR0PS; M10.003; -  
DR InterPro; IPR000562; FN\_Type\_II.  
DR InterPro; IPR000585; Hemopexin.  
DR InterPro; IPR001818; Matrxin.  
DR InterPro; IPR000130; Zn\_MTPeptidase.  
DR Pfam; PF00040; fn2; 3.  
DR Pfam; PF00045; hemopexin; 4.  
DR Pfam; PF00413; Peptidase\_M10; 1.  
DR PRINTS; PR00013; FNTYPEII.  
DR PRINTS; PR00138; MATRININ.  
DR ProDom; PD000995; FN\_Type\_II; 3.  
DR SMART; SM00059; FN2; 3.  
DR SMART; SM00120; HX; 4.  
DR SMART; SM00235; ZnMc; 1.  
DR PROSITE; PS00023; FIBRONECTIN\_2; 3.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
KW Hydrolyase; Metalloprotease; Glycoprotein; zinc; zymogen; Calcium;  
KW Collagen degradation; Extracellular matrix; Repeat; Signal.  
FT SIGNAL 1 29  
FT PROPEP 30 109  
FT CHAIN 110 662  
FT DOMAIN 110 221  
FT DOMAIN 222 396  
FT DOMAIN 397 467  
FT DOMAIN 286 383  
FT DOMAIN 284 341  
FT DOMAIN 342 399  
FT DOMAIN 468 662  
FT SITE 102 102  
FT METAL 403 403  
FT ACT\_SITE 404 404  
FT METAL 407 407  
FT METAL 413 413

FT CARBOHYD 575 575 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 644 644 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT DISULFID 471 662 BY SIMILARITY.  
FT CONFLICT 42 A -> S (IN REF. 2).  
FT CONFLICT 286 A -> G (IN REF. 2).  
FT CONFLICT 369 N -> S (IN REF. 2).  
FT CONFLICT 435 H -> N (IN REF. 2).  
FT CONFLICT 586 A -> S (IN REF. 2).  
SQ SEQUENCE 662 AA; 7496B34B0A21884B CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 662;  
Best Local Similarity 75.0%; Pred. NO. 3.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 KQOQFGL 10  
DB 76 KMQKFFGL 83  
| : |||||  
| : |||||  
  
RESULT 116  
MM02\_CHICK STANDARD; PRT; 663 AA.  
ID MM02\_CHICK  
AC Q90611;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE 72 KDA TYPE IV COLLAGENASE PRECURSOR (EC 3.4.24.24) (72 KDA  
GELATINASE) (MATRIX METALLOPROTEINASE-2) (MMP-2) (GELATINASE A).  
GN MMP2.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OC NCBI\_TaxID=9031;  
OX [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Embryo;  
RX MEDLINE=94280397; PubMed=8010954;  
RA Almes R.T., French D.L., Quigley J.P.;  
RT "Cloning of a 72 kDa matrix metalloproteinase (gelatinase) from  
chicken embryo fibroblasts using gene family PCR: expression of the  
gelatinase increases upon malignant transformation.";  
RL Biochem. J. 300:729-736(1994).  
RN [2]  
RP SEQUENCE OF 27-41 AND 107-122.  
RX MEDLINE=91161603; PubMed=1848240;  
RA Chen J.-M., Almes R.T., Ward G.R., Youngleib G.L., Quigley J.P.;  
RT "Isolation and characterization of a 70-kDa metalloprotease  
(gelatinase) that is elevated in Rous sarcoma virus-transformed  
chicken embryo fibroblasts.";  
RL J. Biol. Chem. 266:5113-5121(1991).  
CC -1- CATALYTIC ACTIVITY: CLEAVAGE OF GELATIN TYPE I AND COLLAGEN TYPES  
IV, V, VII, X. CLEAVES THE COLLAGEN-LIKE SEQUENCE PRO-GLN-GLY-|-  
ILE-ALA-GLY-GLN.  
CC -1- COFACTOR: REQUIRES CALCIUM AND ZINC FOR ACTIVITY.  
CC -1- SUBUNIT: LIGAND FOR INTEGRIN ALPHA-V/BETA-3.  
CC -1- TISSUE SPECIFICITY: PRODUCED BY NORMAL SKIN FIBROBLASTS.  
CC -1- SIMILARITY: CONTAINS 1 HEMOPEXIN-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE II-LIKE DOMAINS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M10A (ZINC  
METALLOPROTEASE) ALSO KNOWN AS MATRININ SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U07775; AAA19596.1; -  
DR HSSP; P08253; IRTG.

```
DR MEROPS; M10.003; --
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Matrixin.
DR InterPro; IPR000130; Zn_MTPeptdse.
DR Pfam; PF00040; fn2; 3.
DR Pfam; PF00045; hemopexin; 4.
DR Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PR00013; FNTYPEII.
DR PRINTS; PR00138; MATRXIN.
DR ProDom; PD000995; FN_Type_II; 3.
DR SMART; SM00059; FN2; 3.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZnMG; 1.
DR PROSITE; PS00023; FIBRONECTIN_2; 3.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
KW Hydrolase; Metalloprotease; Zinc; zymogen; Calcium;
Collagen degradation; Extracellular matrix; Signal.
FT SIGNAL 1 26
FT PROPEP 27 106 ACTIVATION PEPTIDE.
FT CHAIN 107 663 72 KDA TYPE IV COLLAGENASE.
FT DOMAIN 107 218 COLLAGENASE-LIKE.
FT DOMAIN 219 393 COLLAGEN-BINDING.
FT DOMAIN 394 468 COLLAGENASE-LIKE.
FT DOMAIN 223 280 FIBRONECTIN TYPE-II 1.
FT DOMAIN 281 338 FIBRONECTIN TYPE-II 2.
FT DOMAIN 339 396 FIBRONECTIN TYPE-II 3.
FT DOMAIN 469 663 HEMOPEXIN-LIKE.
FT SITE 99 99 CYSTEINE SWITCH (POTENTIAL).
FT METAL 400 400 ZINC (CATALYTIC) (BY SIMILARITY).
FT ACT_SITE 401 401 BY SIMILARITY.
FT METAL 404 404 ZINC (CATALYTIC) (BY SIMILARITY).
FT METAL 410 410 ZINC (CATALYTIC) (BY SIMILARITY).
FT DISULFID 472 663 BY SIMILARITY.
FT CONFLICT 40 40 P -> Q (IN REF. 2).
FT CONFLICT 116 116 W -> T (IN REF. 2).
FT CONFLICT 122 122 T -> I (IN REF. 2).
SQ SEQUENCE 663 AA; 74941 MW; 8D6FDA4E67C3EBCA CRC64;

Query Match 50.8%; Score 31; DB 1; Length 663;
Best Local Similarity 75.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KQOQFFGL 10
DB 73 KMQKFFGL 80

RESULT 117
SNWA_DICDI.
ID SNWA_DICDI STANDARD; PRT; 685 AA.
AC P54705;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SNWA PROTEIN.
OS Dictyostellium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
OX NCBI_TaxID=44689;
RN [1]
RP MEDLINE=97128797; PubMed=8973337;
RA Folk P., Puta F., Krpejskova L., Blahuskova A., Markos A.,
RA Rabino M., Dotti R.P.;
RT "The homolog of chromatin binding protein Bx42 identified in
RT Dictyostellium.";
RL Gene 181:229-231(1996).
CC -!- SIMILARITY: BELONGS TO THE SNW FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U43887; AAB40497.1; --
DR DictyDb; DD00074; SNWA.
FT DOMAIN 31 41 POLY-GLN.
FT DOMAIN 194 360 SNW.
FT DOMAIN 245 253 PRO-RICH.
FT DOMAIN 409 415 POLY-ASP.
FT DOMAIN 539 616 SH2-LIKE DOMAIN.
SQ SEQUENCE 685 AA; 78529 MW; 1DC8521E9997A583 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 685;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPOQ 6
DB 27 KRPPOQ 32

RESULT 118
AKA8_RAT
ID AKA8_RAT STANDARD; PRT; 687 AA.
AC Q63014;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE A-KINASE ANCHOR PROTEIN 8 (A-KINASE ANCHOR PROTEIN 95 KDA) (AKAP 95).
GN AKAP8 OR AKAP95.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pituitary;
RX MEDLINE=94171800; PubMed=8125992;
RA Coghlan V.M., Langeberg L.K., Fernandez A., Lamb N.J., Scott J.D.;
RT "Cloning and characterization of AKAP 95, a nuclear protein that
RT associates with the regulatory subunit of type II CAMP-dependent
RT protein kinase.";
RL J. Biol. Chem. 269:7658-7665(1994).
CC -!- FUNCTION: ANCHORING PROTEIN THAT MEDIATES THE SUBCELLULAR
CC COMPARTMENTATION OF CAMP-DEPENDENT PROTEIN KINASE (PKA TYPE II).
CC -!- SUBUNIT: BINDS TO DIMERIC RII-ALPHA REGULATORY SUBUNIT OF PKA
CC DURING MITOSIS.
CC -!- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
CC REDISTRIBUTED AND DETACHED FROM CONDENSED CHROMATIN DURING MITOSIS
CC (BY SIMILARITY).
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE PROTEIN HAS BEEN
CC DETECTED IN LIVER, FIBROBLASTS, GRANULOSA, MYOBLAST, LYMPHOMA AND
CC SERTOLI CELLS.
CC -!- PTM: PHOSPHORYLATED BY PKC.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U01914; AAA95941.1; --
DR InterPro; IPR000822; Znf_C2H2.
DR SMART; SM00355; Znf_C2H2; 1.
KW Nuclear protein; Zinc-finger; DNA-binding; Phosphorylation.
```

FT ZN\_FING 390 412 C2H2-TYPE.  
FT ZN\_FING 479 502 C2H2-TYPE.  
FT DOMAIN 366 375 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT DOMAIN 568 585 RII-BINDING.  
FT DOMAIN 654 657 POLY-ALA.  
SQ SEQUENCE 687 AA; 76161 MW; 7535F30F18F1B8CB CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 687;  
Best Local Similarity 50.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKPQOFFGL 10  
:|||||  
Db 456 KPKPDPFKGI 465  
  
RESULT 119  
AKA8\_HUMAN STANDARD; PRT; 692 AA.  
AC O43823;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE A-KINASE ANCHOR PROTEIN 8 (A-KINASE ANCHOR PROTEIN 95 KDA) (AKAP 95).  
GN AKAP8 OR AKAP95.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA TISSUE-Cerebellum, and Testis;  
RX MEDLINE=98142017; PubMed=9473338;  
RA Elde T., Coghlan V., Oerstavik S., Holsve C., Solberg R.,  
RA Skaelhegg B.S., Lamb N.J.C., Langeberg L., Fernandez A., Scott J.D.,  
RA Jahnsen T., Tasken K.;  
RT "Molecular cloning, chromosomal localization, and cell cycle-dependent  
RT subcellular distribution of the A-kinase anchoring protein, AKAP95";  
RL Exp. Cell Res. 238:305-316(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V.,  
RA Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwagen S.,  
RA Phan H., Vellasco N., Do L., Regala W., Terry A., Ganes J.,  
RA Danganan L., Erler A., Christensen M., Georgescu A., Avila J., Liu S.,  
RA Attix C., Andreise T., Trankheim M., Amico-Keller G., Coefield J.,  
RA Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Krommiller B.,  
RA Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A.,  
RA Olsen A.S., Carrano A.V.;  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: ANCHORING PROTEIN THAT MEDIATES THE SUBCELLULAR  
CC COMPARTMENTATION OF CAMP-DEPENDENT PROTEIN KINASE (PKA TYPE II).  
CC -1- SUBUNIT: BINDS TO DIMERIC RII-ALPHA REGULATORY SUBUNIT OF PKA  
CC DURING MITOSIS.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.  
CC REDISTRIBUTED AND DETACHED FROM CONDENSED CHROMATIN DURING  
CC MITOSIS.  
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN HEART, LIVER, SKELETAL  
CC MUSCLE, KIDNEY AND PANCREAS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; Y11997; CAA72722.1;  
CC EMBL; AC005785; AAC62838.1;  
CC MM; 604692;  
CC InterPro; IPR000822; Znf-C2H2.

DR SMART; SM00355; Znf\_C2H2; 1.  
KW Nuclear protein; Zinc-finger; DNA-binding.  
FT DOMAIN 107 118 POLY-GLY.  
FT DOMAIN 368 377 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT ZN\_FING 392 414 C2H2-TYPE.  
FT ZN\_FING 481 504 C2H2-TYPE.  
FT DOMAIN 572 589 RII-BINDING (BY SIMILARITY).  
SQ SEQUENCE 692 AA; 76108 MW; CBGD5F014FD94B66 CRC64;  
  
Query Match 50.8%; Score 31; DB 1; Length 692;  
Best Local Similarity 50.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKPQOFFGL 10  
:|||||  
Db 458 KPKPDPFKGI 467  
  
RESULT 120  
PNP\_BACSU STANDARD; PRT; 704 AA.  
AC P50849;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE POLYBONUCLEOTIDE NUCLEOTIDYLTRANSFERASE (EC 2.7.7.8) (POLYNUCLEOTIDE  
DE PHOSPHORYLASE) (PNPASE) (VEGETATIVE PROTEIN 15) (VEG15).  
GN PNP OR COMR.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96423178; PubMed=8825779;  
RA Luttinger A., Hahn J., Dubnau D.;  
RT "Polynucleotide phosphorylase is necessary for competence development  
RT in Bacillus subtilis";  
RL Mol. Microbiol. 19:343-356(1996).  
RN [2]  
RP SEQUENCE OF 1-8 FROM N.A.  
RC STRAIN=168;  
RA Coquard D., Huecas M., Ott M., van Dijk J., van Loon A., Hohmann H.;  
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 1-20.  
RC STRAIN=IS58;  
RX MEDLINE=97443988; PubMed=9298659;  
RA Antelmann H., Bernhardt J., Schmid R., Mach H., Voelker U.,  
RA Hecker M.;  
RT "First steps from a two-dimensional protein index towards a response-  
RT regulation map for Bacillus subtilis";  
RL Electrophoresis 18:1451-1463(1997)  
CC -1- FUNCTION: INVOLVED IN MRNA DEGRADATION. HYDROLYSES SINGLE-STRANDED  
CC POLYBONUCLEOTIDES PROGRESSIVELY IN THE 3' TO 5' DIRECTION. MAY BE  
CC NECESSARY FOR COMPETENCE DEVELOPMENT IN BACILLUS SUBTILIS. MAY BE  
CC NECESSARY FOR MODIFICATION OF THE SREA TRANSCRIPT (STABILIZATION  
CC OR TRANSLATION ACTIVATION).  
CC -1- CATALYTIC ACTIVITY: RNA(N+1) + ORTHOPHOSPHATE = RNA(N) + A  
CC NUCLEOSIDE DIPHOSPHATE.  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- SIMILARITY: CONTAINS 1 KH DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 'S1 MOTIF' DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC



```
DR EMBL; U29668; AAC43595.1; -.
DR EMBL; Z80835; CAB02561.1; -.
DR EMBL; Z99112; CAB13542.1; -.
DR HSSP; P05055; 1SRO.
DR Subtilist; BG11491; pnpA.
DR InterPro; IPR001247; 3_ExoNase.
DR InterPro; IPR000958; KH.
DR InterPro; IPR003029; S1.
DR Pfam; PF00013; KH-domain; 1.
DR Pfam; PF01138; RNase_PH; 2.
DR Pfam; PF00575; S1; 1.
DR SMART; SM00322; KH; 1.
DR SMART; SM00316; S1; 1.
DR PROSITE; PS0084; KH_TYPE_1; 1.
KW Transferase; Nucleotidyltransferase; RNA-binding; Complete proteome.
FT INIT_MET 0
FT DOMAIN 553 612 KH.
FT DOMAIN 622 690 S1 MOTIF.
SQ SEQUENCE 704 AA; 77332 MW; 0E305B6B9B0A7B07 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 704;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQOFFGL 10
||| |||
DB 50 PKPLDFEPL 58

RESULT 121
LIPS.RAT STANDARD; PRT; 768 AA.
ID LIPS.RAT STANDARD; PRT; 768 AA.
AC P15304;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE HORMONE SENSITIVE LIPASE (EC 3.1.1.-) (HSL).
GN LIPE.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=adipose tissue;
RX MEDLINE=89041594; PubMed=3186461;
RA Holm C., Kirchgessner T.G., Svenson K.L., Lusis A.J., Belfrage P.,
RA Schotz M.C.;
RT "Nucleotide sequence of rat adipose hormone sensitive lipase cDNA.";
RL Nucleic Acids Res. 16:9879-9879(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88336885; PubMed=3420405;
RA Holm C., Kirchgessner T.G., Svenson K.L., Fredrikson G., Nilsson S.,
RA Miller C.G., Shively J.E., Heinemann C., Sparkes R.S., Mohandas T.,
RA Lusis A.J., Belfrage P., Schotz M.C.;
RT "Hormone-sensitive lipase: sequence, expression, and chromosomal
RT localization to 19 cent-q13.3.";
RL Science 241:1503-1506(1988).
RN [3]
RP REVISIONS TO 542-555 AND 746-768.
RC TISSUE=adipose tissue;
RA Holm C.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: IN ADIPOSE TISSUE AND HEART, IT PRIMARILY HYDROLYZES
CC STORED TRIGLYCERIDES TO FREE FATTY ACIDS, WHILE IN STEROIDOGENIC
CC TISSUES, IT PRINCIPALLY CONVERTS CHOLESTERYL ESTERS TO FREE
CC CHOLESTEROL FOR STEROID HORMONE PRODUCTION.
CC -!- ENZYME REGULATION: RAPIDLY ACTIVATED BY CAMP-DEPENDENT
CC PHOSPHORYLATION UNDER THE INFLUENCE OF CATECHOLAMINES.
CC DEPHOSPHORYLATION AND INACTIVATION ARE CONTROLLED BY INSULIN.
CC -!- PATHWAY: HORMONE SENSITIVE LIPASE CATALYZES THE RATE LIMITING
```

---

```
CC STEP IN TRIGLYCERIDE LIPOLYSIS.
CC -!- SIMILARITY: BELONGS TO THE "GDXG" FAMILY OF LIPOLYTIC ENZYMES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X51415; CAA35777.1; -.
DR PIR; S03672; LIRTH.
DR InterPro; IPR000379; Est_lip_thioest_actsite.
DR InterPro; IPR002168; Lipolytic_enzyme.
DR PROSITE; PS01173; LIPASE_GDXG_HIS; 1.
DR PROSITE; PS01174; LIPASE_GDXG_SER; 1.
KW Hydrolase; Lipid degradation; Phosphorylation.
FT ACT_SITE 349 349 POTENTIAL.
FT ACT_SITE 423 423 POTENTIAL.
FT MOD_RES 563 563 PHOSPHORYLATION (BY CAPK)
FT MOD_RES 565 565 (BY SIMILARITY).
FT MOD_RES 565 565 PHOSPHORYLATION (BY AMPK)
FT MOD_RES 565 565 (BY SIMILARITY).
SQ SEQUENCE 768 AA; 84169 MW; 90A1F0432DAC8B4C CRC64;

Query Match 50.8%; Score 31; DB 1; Length 768;
Best Local Similarity 83.3%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQ 6
||| |||
DB 332 RRPQQ 337

RESULT 122
E78A.DROME STANDARD; PRT; 864 AA.
ID E78A.DROME STANDARD; PRT; 864 AA.
AC P45447;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ECDYSONE-INDUCIBLE PROTEIN E78-A (DR-78).
GN EIP78C OR E78A OR NR1E1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CANTON-S;
RX MEDLINE=94006562; PubMed=8402914;
RA Stone B.L., Thummel C.S.;
RT "The Drosophila 78C early late puff contains E78, an ecdysone-
RT inducible gene that encodes a novel member of the nuclear hormone
RT receptor superfamily.";
RL Cell 75:307-320(1993).
RN [2]
RP SEQUENCE OF 321-433 FROM N.A.
RX MEDLINE=94060116; PubMed=8241283;
RA Martin-Blanco E., Kornberg T.B.;
RT "DR-78, a novel Drosophila melanogaster genomic DNA fragment highly
RT homologous to the DNA-binding domain of thyroid hormone-retinoid
RT acid-vitamin D receptor subfamily.";
RL Biochim. Biophys. Acta 1216:339-341(1993).
CC -!- FUNCTION: INDUCES THE EARLY LATE PUFF 78C WHICH TRIGGERS PUPARIUM
CC FORMATION AND DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -!- ALTERNATIVE PRODUCTS: THE DIFFERENT FORMS OF PROTEIN E78 ARE
CC PROBABLY PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
CC -!- DEVELOPMENTAL STAGE: THE LONGER FORM, E78A, IS EXPRESSED ONLY
```

CC IN MID-PUPAL STAGES, WHILE THE SHORTER FORM, E78B, IS MAXIMALLY  
CC EXPRESSED IN NEWLY FORMED PREPUFAE.  
CC -1- INDUCTION: BOTH FORMS REQUIRE ECDYSONE FOR ACTIVITY. E78B ALSO  
CC REQUIRES ECDYSONE-INDUCED PROTEINS FOR MAXIMAL EXPRESSION.  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
CC NRI SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U01087; AAA19975.1; -  
DR EMBL; U01088; AAA19976.2; -  
DR EMBL; X73045; CAA51523.1; -  
DR HSP; P03372; IHQ.  
DR FlyBase; FBgn004865; Eip78C.  
DR InterPro; IPR000536; Hormone\_rec\_lig.  
DR InterPro; IPR001723; Strdhormone\_receptor.  
DR InterPro; IPR001628; zf-C4.  
DR Pfam; PF00104; hormone\_rec; 1.  
DR Pfam; PF00105; zf-C4; 1.  
DR PRINTS; PR00047; STROIDFINGER.  
DR PRINTS; PR00398; STRODHORMONER.  
DR SMART; SM00430; HOLI; 1.  
DR SMART; SM00399; ZnF\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
KW Transcription regulation; DNA-binding; Nuclear protein;  
KW Zinc-finger; Alternative splicing.  
FT DNA\_BIND 367 432 NUCLEAR RECEPTOR-TYPE.  
FT ZN\_FING 367 387 C4-TYPE.  
FT ZN\_FING 403 427 C4-TYPE.  
FT DOMAIN 64 80 POLY-GLU.  
FT DOMAIN 182 188 POLY-GLN.  
FT DOMAIN 192 202 POLY-GLN.  
FT DOMAIN 240 247 POLY-SER.  
FT DOMAIN 271 279 POLY-SER.  
FT DOMAIN 312 315 POLY-GLN.  
FT DOMAIN 321 333 POLY-GLN.  
FT DOMAIN 336 339 POLY-GLN.  
FT DOMAIN 346 349 POLY-SER.  
FT DOMAIN 354 357 POLY-ASN.  
FT DOMAIN 481 486 POLY-GLN.  
FT DOMAIN 490 500 POLY-GLN.  
FT DOMAIN 546 554 POLY-ASN.  
FT VARSPIC 1 474 MISSING (IN TRUNCATED E78B ISOFORM).  
FT CONFLICT 321 331 OLQOOOQHQQ -> SCNSSSTSSR (IN REF. 2).  
FT CONFLICT 430 433 AGWS -> VGKM (IN REF. 2).  
SQ SEQUENCE 864 AA; 95865 MW; 1A5CE6C39F31E994 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 864;  
Best Local Similarity 66.7%; Pred. No. 4e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQQFFGL 10  
Db 335 PQQQSFGL 343  
1: || |||

RESULT 123  
YB1C\_SCHPO STANDARD; PRT; 922 AA.  
AC PB7177;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HYPOTHETICAL 103.4 KDA TRP-ASP REPEATS CONTAINING PROTEIN C3D6.12 IN  
DE CHROMOSOME II.  
GN SPC3D6.12.

OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA Hilbert H., Duesterhoeft A., Wood V., Rajandream M.A., Barrell B.G.;  
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO YEAST DIP2.  
CC -1- SIMILARITY: CONTAINS 13 WD REPEATS (TRP-ASP DOMAINS).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; Z95620; CAB09121.1; -  
DR InterPro; IPR001680; WD40.  
DR Pfam; PF00400; WD40; 11.  
DR PRINTS; PR00320; GPROTEINBRPT.  
DR SMART; SM00320; WD40; 11.  
DR PROSITE; PS00678; WD\_REPEATS\_1; FALSE\_NEG.  
DR PROSITE; PS50082; WD\_REPEATS\_2; 7.  
DR PROSITE; PS50294; WD\_REPEATS\_REGION; 1.  
KW Hypothetical protein; Repeat; WD repeat.  
FT REPEAT 63 101 WD 1.  
FT REPEAT 102 141 WD 2.  
FT REPEAT 169 210 WD 3.  
FT REPEAT 213 252 WD 4.  
FT REPEAT 268 307 WD 5.  
FT REPEAT 336 377 WD 6.  
FT REPEAT 394 432 WD 7.  
FT REPEAT 475 514 WD 8.  
FT REPEAT 532 571 WD 9.  
FT REPEAT 574 615 WD 10.  
FT REPEAT 616 655 WD 11.  
FT REPEAT 658 695 WD 12.  
FT REPEAT 695 713 WD 13.  
SQ SEQUENCE 922 AA; 103427 MW; 578C01A1D71162C9 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 922;  
Best Local Similarity 45.5%; Pred. No. 4.3e+02;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQOFFGLM 11  
Db 7 RYEPTEFGVI 17  
1: -1: |||;

RESULT 124  
PMAB\_ARATH STANDARD; PRT; 956 AA.  
AC Q9LV11;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ATPASE 11, PLASMA MEMBRANE-TYPE (EC 3.6.3.6) (PROTON PUMP 11).  
GN AH11 OR AT5G62670 OR MRC21.9.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RX MEDLINE=20181125; PubMed=10718197;

RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,  
RA Tabata S.,  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence  
RT features of the regions of 3,076,755 bp covered by sixty PI and TAC  
RT clones";  
RL DNA Res. 7:31-63(2000).  
CC -!- FUNCTION: THE PLASMA MEMBRANE H+ ATPASE OF PLANTS AND FUNGI  
CC GENERATES A PROTON GRADIENT THAT DRIVES THE ACTIVE TRANSPORT OF  
CC NUTRIENTS BY H+-SYMPORT. THE RESULTING EXTERNAL ACIDIFICATION  
CC AND/OR INTERNAL ALKALIZATION MAY MEDIATE GROWTH RESPONSES (BY  
CC SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(IN) -> ADP + PHOSPHATE +  
CC H(+)(OUT).  
CC -!- SUBUNIT: BINDS TO 14-3-3 PROTEINS. THE BINDING IS INDUCED BY  
CC PHOSPHORYLATION OF THR-955. BINDING TO 14-3-3 PROTEINS ACTIVATES  
CC THE H+-ATPASE (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
CC -!- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
CC (E1-E2 ATPASES). SUBFAMILY IIIA.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AB020751; BAA97214.1; -.  
DR InterPro; IPR001757; E1-E2\_ATPase.  
DR InterPro; IPR000695; HATPase.  
DR InterPro; IPR001454; Hydrolase.  
DR Pfam; PF00122; E1-E2\_ATPase; 1.  
DR Pfam; PF00702; Hydrolase; 1.  
DR PRINTS; PR00120; HATPase.  
DR PROSITE; PS00154; ATPASE\_E1\_E2; 1.  
KW Hydrolase; Hydrogen ion transport; Transmembrane; Phosphorylation;  
KW ATP-binding; Metal-binding; Magnesium; Multigene family.  
FT DOMAIN 1 65  
FT TRANSMEM 66 85  
FT DOMAIN 86 97  
FT TRANSMEM 98 118  
FT DOMAIN 119 247  
FT TRANSMEM 248 268  
FT DOMAIN 269 277  
FT TRANSMEM 278 295  
FT DOMAIN 296 647  
FT TRANSMEM 648 669  
FT DOMAIN 670 674  
FT TRANSMEM 675 697  
FT DOMAIN 698 713  
FT TRANSMEM 714 734  
FT DOMAIN 735 759  
FT TRANSMEM 760 780  
FT DOMAIN 781 792  
FT TRANSMEM 793 813  
FT DOMAIN 814 821  
FT TRANSMEM 822 842  
FT DOMAIN 843 956  
FT MOD\_RES 333 333  
FT MOD\_RES 955 955  
FT METAL 592 592  
FT METAL 596 596  
FT SITE 954 956  
FT SEQUENCE 956 AA; 105122 MW; CA59212B16B9C5BD CRC64;  
Query Match 50.8%; Score 31; DB 1; Length 956;  
Best Local Similarity 75.0%; Pred. No. 4.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 4 PQQFFGLM 11

Db 480 PQQFFGLM 487

| | | | |

## RESULT 125

## ATS9 HUMAN

ID ATS9 HUMAN STANDARD; PRT; 1629 AA.  
AC O9P2N4; Q9NR29;  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE ADAM-TS 9 PRECURSOR (EC 3.4.24.-) (A DISINTEGRIN AND METALLOPROTEINASE  
DE WITH THROMBOSPONDIN MOTIFS 9) (ADAMTS-9) (ADAM-TS9).  
GN ADAMTS9 OR KIAA1312.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eumalia; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (SHORT ISOFORM).  
RC TISSUE=Fetal;  
RX MEDLINE=20396138; PubMed=10936055;  
RA Clark M.E., Kelnar G.S., Turbeville L.A., Boyer A., Arden K.A.,  
RA Maki R.A.;  
RT "ADAMTS 9, a novel member of the ADAM-TS/Metallopondin gene  
RT family";  
RL Genomics 67:343-350(2000).  
RN [2]  
RP SEQUENCE OF 159-1629 FROM N.A. (LONG ISOFORM).  
RC TISSUE=Brain;  
RX MEDLINE=20181126; PubMed=10718198;  
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirose M., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XVI.  
RT The complete sequences of 150 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 7:65-73(2000).  
CC -!- COFACTOR: BINDS ONE ZINC ION (BY SIMILARITY).  
CC -!- SUBCELLULAR LOCATION: SECRETED. ASSOCIATED WITH THE EXTRACELLULAR  
CC MATRIX (BY SIMILARITY).  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN ALL FETAL TISSUES.  
CC EXPRESSED SLIGHTLY IN ADULT OVARY, PANCREAS, HEART, KIDNEY, LUNG,  
CC PLACENTA. ALSO DETECTED IN SPINAL CORD AND BRAIN. NOT DETECTED IN  
CC COLON, SMALL INTESTINE, TESTIS, LIVER, SKELETAL MUSCLE, SPLEEN OR  
CC THYMUS.  
CC -!- DOMAIN: THE SPACER DOMAIN AND THE TSP TYPE 1 DOMAINS ARE IMPORTANT  
CC FOR A TIGHT INTERACTION WITH THE EXTRACELLULAR MATRIX (BY  
CC SIMILARITY).  
CC -!- PTM: THE PRECURSOR IS CLEAVED BY A URIN ENDOPEPTIDASE (BY  
CC SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M12B (ZINC  
CC METALLOPROTEASE); ALSO KNOWN AS THE REPROLYSIN SUBFAMILY.  
CC -!- SIMILARITY: CONTAINS 1 DISINTEGRIN-LIKE DOMAIN.  
CC -!- SIMILARITY: CONTAINS 11 TSP TYPE-1 DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF261918; AAF89106.1; -.  
DR EMBL; AB037733; BAA92550.1; -.  
DR MM; 605421; -.  
DR InterPro; IPR001590; Reprolysin.  
DR InterPro; IPR000884; TSPI.  
DR InterPro; IPR000130; Zn\_MTPetdse.  
DR Pfam; PF01421; Reprolysin; 1.  
DR Pfam; PF00090; tsp.1; 11.  
DR SMART; SM00209; TSP1; 13.

```
DR PROSITE; PS00215; ADAM_MEPRO; 1.
DR PROSITE; PS00427; DISINTEGRINS; FALSE_NEG.
DR PROSITE; PS00092; TSP1; 9.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Hydrolase; Metalloprotease; zinc; Signal; Glycoprotein; Zymogen;
FT SIGNAL 1 18
FT PROPEP 19 287
FT CHAIN 288 1629
FT DOMAIN 509 587
FT DOMAIN 589 642
FT DOMAIN 645 752
FT DOMAIN 753 880
FT DOMAIN 999 1053
FT DOMAIN 1056 1108
FT DOMAIN 1111 1156
FT DOMAIN 1184 1239
FT DOMAIN 1240 1295
FT DOMAIN 1332 1383
FT DOMAIN 1386 1439
FT DOMAIN 1445 1498
FT DOMAIN 1501 1554
FT DOMAIN 1562 1612
FT DOMAIN 88 96
FT SITE 223 223
FT METAL 434 434
FT ACT_SITE 435 435
FT METAL 438 438
FT METAL 444 444
FT CARBOHYD 112 112
FT CARBOHYD 135 135
FT CARBOHYD 171 171
FT CARBOHYD 231 231
FT CARBOHYD 749 749
FT CARBOHYD 840 840
FT CARBOHYD 1213 1213
FT CARBOHYD 1267 1267
FT VARSPPLIC 1064 1064
FT VARSPPLIC 1073 1629
FT - CONFLICT 367 367
FT - CONFLICT F -> L (IN REF. 1).
SQ SEQUENCE 1629 AA; 182649 MW; 182649 MW; C1C4CEFF58B8941F CRC64;

Query Match 50.8%; Score 31; DB 1; Length 1629;
Best Local Similarity 83.3%; Pred. No. 7.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 KPQQFF 8
DB 875 KPQQFY 880

RESULT 126
ACVS_CEPAC STANDARD; PRT; 3712 AA.
AC P25464;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE DELTA-(L-ALPHA-AMINOADIPYL)-L-CYSTEINYL-D-VALINE SYNTHETASE
DE (EC 6.-.-.-) (ACV SYNTHETASE) (ACVS).
GN PCBAB.
OS Cephalosporium acremonium (Acremonium chrysogenum).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Hypocreales; Hypocreaceae; mitosporic Hypocreaceae; Acremonium.
OX NCBI_TaxID=5044;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91177827; PubMed=1706706;
RA Gutierrez S., Diez B., Montenegro E., Martin J.F.;
RT "Characterization of the Cephalosporium acremonium pcbab gene
RT encoding alpha-aminoadipyl-cysteine-yl-valine synthetase, a large
RT multidomain peptide synthetase: linkage to the pcbC gene as a cluster
```

```
RT of early cephalosporin biosynthetic genes and evidence of multiple
RT functional domains."
RL J. Bacteriol. 173:2354-2365(1991).
RN [2]
RP PARTIAL SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=ATCC 11550;
RX MEDLINE=91168300; PubMed=2076552;
RA Hoskins J.A., O'Callaghan N., Queener S.W., Cantwell C.A., Wood J.S.,
RA Chen V.J., Skatrud P.L.;
RT "Gene disruption of the pcbAB gene encoding ACV synthetase in
RT Cephalosporium acremonium."
RL Curr. Genet. 18:523-530(1990).
CC -|- FUNCTION: EACH OF THE CONSTITUENT AMINO ACIDS OF THE TRIPETIDE
CC ACV ARE ACTIVATED AS AMINOACYL-ADENYLATES WITH PEPTIDE BONDS
CC FORMED THROUGH THE PARTICIPATION OF AMINO ACID THIOLESTER
CC INTERMEDIATES.
CC -|- COFACTOR: CONTAINS 3 COVALENTLY BOUND PHOSPHOPANTETHEINES
CC (POTENTIAL).
CC -|- PATHWAY: FIRST STEP IN THE BIOSYNTHESIS OF PENICILLIN AND
CC CEPHALOSPORIN.
CC -|- SIMILARITY: BELONGS TO THE ATP-DEPENDENT AMP-BINDING ENZYME
CC FAMILY.
DR PIR; A38531; YGCEVC.
DR HSP; P14687; IAMU.
DR InterPro; IPR000873; AMP-bind.
DR InterPro; IPR001242; DUF4.
DR InterPro; IPR000379; Est_lip_thioest_actsite.
DR InterPro; IPR003880; Phosphopant_attach.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00501; AMP-binding; 3.
DR Pfam; PF00668; Condensation; 3.
DR Pfam; PF00550; pp-binding; 3.
DR Pfam; PF00975; Thioesterase; 1.
DR PRINTS; PR00154; AMPBINDING.
DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
DR PROSITE; PS00455; AMP_BINDING; 3.
DR PROSITE; PS00075; ACP_DOMAIN; 3.
KW Ligase; Antibiotic biosynthesis; Multifunctional enzyme;
KW Repeat; Phosphopantetheine.
FT REPEAT 234 1062 DOMAIN 1 (ADIPATE-ACTIVATING).
FT REPEAT 1335 2162 DOMAIN 2 (CYSTEINE-ACTIVATING).
FT REPEAT 2409 3387 DOMAIN 3 (VALINE-ACTIVATING).
FT DOMAIN 795 864 ACYL CARRIER (ACP) 1.
FT DOMAIN 1880 1953 ACYL CARRIER (ACP) 2.
FT DOMAIN 2960 3027 ACYL CARRIER (ACP) 3.
FT BINDING 827 827 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT BINDING 1916 1916 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT BINDING 2990 2990 PHOSPHOPANTETHEINE (BY SIMILARITY).
FT ACT_SITE 3568 3568 THIOESTERASE (BY SIMILARITY).
SQ SEQUENCE 3712 AA; 414767 MW; 4EE3C1EB5EBEF9B7 CRC64;

Query Match 50.8%; Score 31; DB 1; Length 3712;
Best Local Similarity 71.4%; Pred. No. 1.7e+03;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOQF 7
DB 2189 RRPRAQF 2195

RESULT 127
ABCL_MOUSE STANDARD; PRT; 2261 AA.
ID ABCL_MOUSE
AC P41233;
DT 01-FEB-1995 (Rel. 31, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE ATP-BINDING CASSETTE, SUB-FAMILY A, MEMBER 1 (ATP-BINDING CASSETTE
DE TRANSPORTER 1) (ATP-BINDING CASSETTE 1) (ABC-1).
GN ABCAL OR ABCL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DBA/2; TISSUE=Macrophage;  
RX MEDLINE=94375008; PubMed=8088782;  
RA Luciani M.F., Denizot F., Savary S., Mattei M.-G., Chimini G.;  
RT "Cloning of two novel ABC transporters mapping on human chromosome  
9.";  
RL Genomics 21:150-159(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J;  
RA Qiu Y., Cavellier L., Chiu S., Rubin E., Cheng J.-F.;  
RT "Human and mouse ABCA1 comparative sequencing and transgenesis studies  
identify potential regulatory sequences.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CAMP-DEPENDENT AND SULFONYLUREA-SENSITIVE ANION  
CC TRANSPORTER. KEY GATEKEEPER INFLUENCING INTRACELLULAR CHOLESTEROL  
CC TRANSPORT (BY SIMILARITY).  
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN ADULT TISSUES. HIGHEST  
CC LEVELS ARE FOUND IN PREGNANT UTERUS AND UTERUS.  
CC -1- DOMAIN: MULTIFUNCTIONAL POLYPEPTIDE WITH TWO HOMOLOGOUS HALVES,  
CC EACH CONTAINING AN HYDROPHOBIC MEMBRANE-ANCHORING DOMAIN AND AN  
CC ATP BINDING CASSETTE (ABC) DOMAIN.  
CC -1- SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY  
CC (ABC TRANSPORTERS). MDR SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
DR EMBL; X75926; CAA53530.1; ALT\_INIT.  
DR EMBL; AF287263; AAG39073.1; ALT\_INIT.  
DR MGD; MGI:99607; Abca1.  
DR InterPro; IPR003593; AAA.  
DR InterPro; IPR003439; ABC\_transportr.  
DR InterPro; IPR001687; ATP\_GTP\_A.  
DR InterPro; IPR003838; DUF214.  
DR InterPro; IPR001865; Ribosomal\_S2.  
DR InterPro; IPR000897; SRP54.  
DR Pfam; PF00005; ABC\_tran; 2.  
DR Pfam; PF02687; DUF214; 1.  
DR Pfam; PF00318; Ribosomal\_S2; 1.  
DR Pfam; PF00448; SRP54; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER; 1.  
KW ATP-binding; Glycoprotein; Transmembrane; Transport.  
FT TRANSMEM . 26 42 POTENTIAL.  
FT TRANSMEM . 640 656 POTENTIAL.  
FT TRANSMEM . 690 706 POTENTIAL.  
FT TRANSMEM . 717 733 POTENTIAL.  
FT TRANSMEM . 749 765 POTENTIAL.  
FT TRANSMEM . 771 787 POTENTIAL.  
FT TRANSMEM . 1041 1057 POTENTIAL.  
FT TRANSMEM . 1351 1367 POTENTIAL.  
FT TRANSMEM . 1661 1677 POTENTIAL.  
FT TRANSMEM . 1708 1724 POTENTIAL.  
FT TRANSMEM . 1737 1753 POTENTIAL.  
FT TRANSMEM . 1775 1791 POTENTIAL.  
FT TRANSMEM . 1854 1870 POTENTIAL.  
FT NP\_BIND . 933 940 ATP (POTENTIAL).  
FT NP\_BIND . 1946 1953 ATP (POTENTIAL).  
FT CARBOHYD . 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD . 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD . 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD . 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD . 196 196 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD . 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 337 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 349 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 400 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 478 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 489 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 521 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 820 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1144 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1453 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1499 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1504 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1637 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2044 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 1567 MISSING (IN REF. 2).  
FT CONFLICT 2024 MISSING (IN REF. 2).  
SQ SEQUENCE 2261 AA; 254011 MW; FA6G2B21FD1D09F9 CRC64;

Query Match 50.0%; Score 30.5; DB 1; Length 2261;  
Best Local Similarity 77.8%; Pred. NO. 1.3e+03;  
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Oy 3 KP-QOFFGL 10  
Db 2153 KPQVEFFGL 2161

Search completed: April 1, 2002, 16:20:17  
Job time: 163 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2002, 16:18:20 ; Search time 38.86 Seconds

(without alignments)  
20.968 Million cell updates/sec

Title: US-09-988-792-2

Perfect score: 71

Sequence: 1 RPKQQWFWM 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 434

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 50%

Maximum Match 100%

Listing first 1000 summaries

Database :

A\_Geneseq\_1101.\*

- 1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.\*
- 2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.\*
- 3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.\*
- 4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.\*
- 5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.\*
- 6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.\*
- 7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.\*
- 8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.\*
- 9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.\*
- 10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.\*
- 11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.\*
- 12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.\*
- 13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.\*
- 14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.\*
- 15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.\*
- 16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.\*
- 17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.\*
- 18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.\*
- 19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.\*
- 20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.\*
- 21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.\*
- 22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description        |
|------------|-------|-------------|--------|----|--------------------|
| 1          | 71    | 100.0       | 11     | 4  | Sequence of peptid |
| 2          | 71    | 100.0       | 11     | 9  | Sequence of neurop |
| 3          | 71    | 100.0       | 11     | 19 | Substance P analog |
| 4          | 71    | 100.0       | 11     | 20 | Human tachykinin a |
| 5          | 71    | 100.0       | 11     | 22 | Human tachykinin a |
| 6          | 71    | 100.0       | 11     | 22 | Chimeric analgesic |
| 7          | 71    | 100.0       | 12     | 22 | Tachykinins peptid |
| 8          | 68    | 95.8        | 11     | 5  | Chimeric analgesic |
| 9          | 68    | 95.8        | 11     | 5  | Substance P analog |
| 10         | 68    | 95.8        | 11     | 9  | Sequence of neurop |
| 11         | 68    | 95.8        | 11     | 11 | Sequence of neurop |
| 12         | 68    | 95.8        | 11     | 11 | D-arginine 1, D-pr |

Substance P analog  
Substance P analog  
Human tachykinin a  
Tachykinins peptid  
Galanin(1-12)-Pro-  
Galanin peptide SE  
Galanin peptide SE  
Spantide analogue,  
Substance P analog  
Substance P analog  
Sequence of neurop  
Chimeric analgesic  
Chimeric analgesic  
Substance P analog  
Substance P analog  
Substance P analog  
Substance P analog  
Amino acid sequenc  
Tachykinins peptid  
Amino acid sequenc  
Sequence of neurop  
Substance P analog  
Bradykinin recepto  
Sequence of peptid  
Substance P [Tyr7]  
Neurokinine 1 liga  
Human tachykinin a  
Sequence of neurop  
Substance P analog  
Substance P analog  
Sequence of undeca  
Sequence of neurop  
Sialic acid-bonded  
Undecapeptide subs  
Substance P [Me-Le  
Substance P [MeMet  
Substance P [Me-Ph  
Substance P [Me-Gl  
Substance P [Me-Gl  
Substance P [p-Chl  
Substance P. Synt  
Neurokinin 1 recep  
Neurokinin 1 recep  
Substance P peptid  
Substance P. Synt  
Mono-DTPA-Arg1 Sub  
Substance P peptid  
Substrate P analog  
Substrate P report  
Non-crosslinked pr  
Substance P. Synt  
Human tachykinin a  
Human tachykinin a  
Human tachykinin a  
Human tachykinin a  
Human tachykinin a  
Human tachykinin a  
Substance P deriva  
Substance P deriva  
Peptide substrate  
Human/rat tachykin  
Cell differentiat  
Amino acid sequenc  
Substance P peptid  
Amino acid sequenc  
Chimeric analgesic

292 886 60

|     |    |      |     |    |          |                     |     |    |      |     |    |          |                      |
|-----|----|------|-----|----|----------|---------------------|-----|----|------|-----|----|----------|----------------------|
| 85  | 48 | 67.6 | 11  | 22 | AAB82070 | Substance P, Unid   | 158 | 43 | 60.6 | 10  | 6  | AAP50633 | Substance P-like P   |
| 86  | 48 | 67.6 | 11  | 22 | AAB91411 | Tachykinins peptid  | 159 | 43 | 60.6 | 10  | 13 | AAR21933 | Human tachykinin a   |
| 87  | 48 | 67.6 | 11  | 22 | AAB91436 | Tachykinins peptid  | 160 | 43 | 60.6 | 10  | 20 | AAW92663 | Tachykinins peptid   |
| 88  | 48 | 67.6 | 11  | 22 | AAB91449 | Tachykinins peptid  | 161 | 43 | 60.6 | 10  | 22 | AAB91423 | Tachykinins peptid   |
| 89  | 48 | 67.6 | 11  | 22 | AAB91450 | Tachykinins peptid  | 162 | 43 | 60.6 | 10  | 22 | AAB91427 | Tachykinins peptid   |
| 90  | 48 | 67.6 | 11  | 22 | AAB50544 | Prolyl endopeptida  | 163 | 43 | 60.6 | 10  | 22 | AAB91445 | Tachykinins peptid   |
| 91  | 48 | 67.6 | 11  | 22 | AAB50306 | Substance P, Unid   | 164 | 43 | 60.6 | 11  | 13 | AAR21936 | Substance P [pro 1   |
| 92  | 48 | 67.6 | 12  | 14 | AAR32798 | Tyr-1 substance P   | 165 | 43 | 60.6 | 11  | 13 | AAR21936 | Substance P or (7-   |
| 93  | 48 | 67.6 | 12  | 16 | AAR85244 | Substance P analog  | 166 | 43 | 60.6 | 11  | 13 | AAR21941 | Substance P [pro 1   |
| 94  | 48 | 67.6 | 12  | 20 | AAV03157 | Substance P-Glycin  | 167 | 43 | 60.6 | 11  | 13 | AAR21944 | Human tachykinin a   |
| 95  | 48 | 67.6 | 12  | 20 | AAW94412 | Cancer protease-se  | 168 | 43 | 60.6 | 11  | 20 | AAW92709 | Human tachykinin a   |
| 96  | 48 | 67.6 | 12  | 22 | AAW92769 | Amino acid sequenc  | 169 | 43 | 60.6 | 11  | 20 | AAW92717 | Human tachykinin a   |
| 97  | 48 | 67.6 | 12  | 22 | AAW92772 | Amino acid sequenc  | 170 | 43 | 60.6 | 11  | 20 | AAW92718 | Human tachykinin a   |
| 98  | 48 | 67.6 | 12  | 22 | AAW92775 | Amino acid sequenc  | 171 | 43 | 60.6 | 11  | 20 | AAW92667 | Human tachykinin a   |
| 99  | 48 | 67.6 | 12  | 22 | AAB84528 | Amino acid sequenc  | 172 | 43 | 60.6 | 11  | 20 | AAW92668 | Human tachykinin a   |
| 100 | 48 | 67.6 | 12  | 22 | AAB98867 | Chimeric analgesic  | 173 | 43 | 60.6 | 11  | 20 | AAW92670 | Human tachykinin a   |
| 101 | 48 | 67.6 | 12  | 22 | AAB98870 | Chimeric analgesic  | 174 | 43 | 60.6 | 11  | 20 | AAW92672 | Human tachykinin a   |
| 102 | 48 | 67.6 | 12  | 22 | AAB98873 | Chimeric analgesic  | 175 | 43 | 60.6 | 11  | 21 | AAB08614 | Peptide identified   |
| 103 | 48 | 67.6 | 13  | 20 | AAV03158 | Substance P-Glycin  | 176 | 43 | 60.6 | 11  | 22 | AAW99350 | Substance P tachy    |
| 104 | 48 | 67.6 | 13  | 22 | AAW92770 | Amino acid sequenc  | 177 | 43 | 60.6 | 898 | 18 | AAW14777 | Granulosis virus 1   |
| 105 | 48 | 67.6 | 13  | 22 | AAW92773 | Amino acid sequenc  | 178 | 43 | 60.6 | 902 | 18 | AAW14285 | H. armigera granu    |
| 106 | 48 | 67.6 | 13  | 22 | AAW92776 | Amino acid sequenc  | 179 | 43 | 60.6 | 7   | 4  | AAP30469 | Sequence of polype   |
| 107 | 48 | 67.6 | 13  | 22 | AAB98868 | Chimeric analgesic  | 180 | 42 | 59.2 | 11  | 13 | AAR21960 | Cyclic substance P   |
| 108 | 48 | 67.6 | 13  | 22 | AAB98871 | Chimeric analgesic  | 181 | 42 | 59.2 | 11  | 13 | AAR21939 | Substance P [ile 8   |
| 109 | 48 | 67.6 | 13  | 22 | AAB98874 | Chimeric analgesic  | 182 | 42 | 59.2 | 11  | 13 | AAR21949 | Substance P [pro 3   |
| 110 | 48 | 67.6 | 14  | 20 | AAV03159 | Substance P-Glycin  | 183 | 42 | 59.2 | 11  | 20 | AAW92683 | Human tachykinin a   |
| 111 | 48 | 67.6 | 14  | 22 | AAW92771 | Amino acid sequenc  | 184 | 42 | 59.2 | 11  | 20 | AAW92669 | Human tachykinin a   |
| 112 | 48 | 67.6 | 14  | 22 | AAW92774 | Amino acid sequenc  | 185 | 42 | 59.2 | 11  | 20 | AAW92673 | Human tachykinin a   |
| 113 | 48 | 67.6 | 14  | 22 | AAW92777 | Amino acid sequenc  | 186 | 42 | 59.2 | 11  | 22 | AAW92665 | Complex sugar bou    |
| 114 | 48 | 67.6 | 14  | 22 | AAW92776 | Chimeric analgesic  | 187 | 42 | 59.2 | 398 | 22 | AAW92665 | Human transporter    |
| 115 | 48 | 67.6 | 14  | 22 | AAW92777 | Chimeric analgesic  | 188 | 41 | 57.7 | 8   | 19 | AAW50976 | Substance P analog   |
| 116 | 48 | 67.6 | 14  | 22 | AAW92778 | Chimeric analgesic  | 189 | 41 | 57.7 | 8   | 20 | AAW92711 | Human tachykinin a   |
| 117 | 48 | 67.6 | 14  | 22 | AAW92779 | Tachykinins peptid  | 190 | 41 | 57.7 | 9   | 13 | AAR21932 | Substance P (1-9)    |
| 118 | 48 | 67.6 | 14  | 22 | AAW92780 | Substance P analog  | 191 | 41 | 57.7 | 9   | 20 | AAV03162 | Substance P [pro 3   |
| 119 | 48 | 67.6 | 14  | 22 | AAW92781 | Human beta-preprot  | 192 | 41 | 57.7 | 9   | 20 | AAW92665 | Human tachykinin a   |
| 120 | 48 | 67.6 | 14  | 22 | AAW92782 | Human atypical tac  | 193 | 41 | 57.7 | 9   | 20 | AAW92665 | Human tachykinin a   |
| 121 | 48 | 67.6 | 14  | 22 | AAW92783 | DB389-SP-Gly fusi   | 194 | 41 | 57.7 | 9   | 22 | AAW92665 | Chimeric analgesic   |
| 122 | 48 | 67.6 | 14  | 22 | AAW92784 | Amyloid precursor   | 195 | 41 | 57.7 | 9   | 22 | AAW92665 | Tachykinins peptid   |
| 123 | 48 | 67.6 | 14  | 22 | AAW92785 | Amyloid precursor   | 196 | 41 | 57.7 | 10  | 22 | AAW92665 | Tachykinins peptid   |
| 124 | 48 | 67.6 | 14  | 22 | AAW92786 | APP-REP 751 protei  | 197 | 41 | 57.7 | 10  | 22 | AAW92665 | Tachykinins peptid   |
| 125 | 48 | 67.6 | 14  | 22 | AAW92787 | Amyloid precursor   | 198 | 41 | 57.7 | 11  | 13 | AAR21940 | Substance P [pro 1   |
| 126 | 48 | 67.6 | 14  | 22 | AAW92788 | APP-REP 751 amyloi  | 199 | 41 | 57.7 | 11  | 20 | AAW92716 | Human tachykinin a   |
| 127 | 48 | 67.6 | 14  | 22 | AAW92789 | Amyloid precursor   | 200 | 41 | 57.7 | 11  | 20 | AAW92721 | Human tachykinin a   |
| 128 | 48 | 67.6 | 14  | 22 | AAW92790 | Amyloid precursor   | 201 | 41 | 57.7 | 11  | 20 | AAW92721 | Substance P analog   |
| 129 | 48 | 67.6 | 14  | 22 | AAW92791 | APP-REP 751 protei  | 202 | 41 | 57.7 | 17  | 21 | AAW92721 | Protein which is s   |
| 130 | 48 | 67.6 | 14  | 22 | AAW92792 | Amyloid precursor   | 203 | 41 | 57.7 | 163 | 20 | AAV37658 | Tachykinins peptid   |
| 131 | 47 | 66.2 | 11  | 13 | AAR21958 | Substance P [Ala 9  | 204 | 40 | 56.3 | 10  | 22 | AAW92721 | Cyclic substance P   |
| 132 | 47 | 66.2 | 11  | 20 | AAW92674 | Human tachykinin a  | 205 | 40 | 56.3 | 11  | 13 | AAR21965 | Substance P [D-Ala   |
| 133 | 47 | 66.2 | 11  | 20 | AAW92675 | Human tachykinin a  | 206 | 40 | 56.3 | 11  | 13 | AAR21964 | Human tachykinin a   |
| 134 | 46 | 64.8 | 11  | 13 | AAR21935 | Substance P [pro 9  | 207 | 40 | 56.3 | 11  | 22 | AAW92682 | Human tachykinin a   |
| 135 | 46 | 64.8 | 11  | 13 | AAR21943 | Substance P [Met 7  | 208 | 40 | 56.3 | 11  | 22 | AAW50311 | Previn peptide #3    |
| 136 | 46 | 64.8 | 11  | 18 | AAW13611 | Spantide II, a sub  | 209 | 40 | 56.3 | 155 | 20 | AAW94272 | Fat-derived eosino   |
| 137 | 46 | 64.8 | 11  | 20 | AAW92677 | Human tachykinin a  | 210 | 40 | 56.3 | 188 | 22 | AAW94272 | Human AFP protein    |
| 138 | 46 | 64.8 | 11  | 20 | AAW92678 | Human tachykinin a  | 211 | 40 | 56.3 | 285 | 21 | AAW94272 | Arabidopsis thalia   |
| 139 | 46 | 64.8 | 11  | 20 | AAW92679 | Human tachykinin a  | 212 | 40 | 56.3 | 285 | 21 | AAW94272 | Arabidopsis thalia   |
| 140 | 46 | 64.8 | 11  | 22 | AAW92680 | Tachykinins peptid  | 213 | 40 | 56.3 | 289 | 21 | AAW94272 | Arabidopsis thalia   |
| 141 | 46 | 64.8 | 11  | 22 | AAW92681 | Tachykinins peptid  | 214 | 40 | 56.3 | 293 | 21 | AAW94272 | Arabidopsis thalia   |
| 142 | 45 | 63.4 | 8   | 19 | AAW50973 | Substance P analog  | 215 | 40 | 56.3 | 293 | 21 | AAW94272 | Arabidopsis thalia   |
| 143 | 45 | 63.4 | 8   | 19 | AAW50975 | Substance P analog  | 216 | 40 | 56.3 | 293 | 21 | AAW94272 | Arabidopsis thalia   |
| 144 | 45 | 63.4 | 11  | 13 | AAR21937 | Substance P or (7-  | 217 | 40 | 56.3 | 293 | 21 | AAW94272 | Arabidopsis thalia   |
| 145 | 45 | 63.4 | 11  | 13 | AAR21951 | Substance P [Glu 3  | 218 | 40 | 56.3 | 317 | 21 | AAW94272 | Arabidopsis thalia   |
| 146 | 45 | 63.4 | 11  | 13 | AAR21952 | Neurokinine 1 liga  | 219 | 40 | 56.3 | 317 | 21 | AAW94272 | Arabidopsis thalia   |
| 147 | 45 | 63.4 | 11  | 14 | AAR21953 | Neurokinine 1 recep | 220 | 40 | 56.3 | 321 | 21 | AAW94272 | Arabidopsis thalia   |
| 148 | 45 | 63.4 | 11  | 16 | AAW09003 | Substance P analog  | 221 | 40 | 56.3 | 363 | 21 | AAW94272 | A prenyltransfer     |
| 149 | 45 | 63.4 | 11  | 18 | AAW33181 | Mono-DTPA-Lys1 Sub  | 222 | 40 | 56.3 | 367 | 21 | AAW94272 | Arabidopsis thalia   |
| 150 | 45 | 63.4 | 11  | 19 | AAW97775 | Substance P, Mamm   | 223 | 40 | 56.3 | 367 | 21 | AAW94272 | Arabidopsis thalia   |
| 151 | 45 | 63.4 | 11  | 20 | AAW92679 | Substance P analog  | 224 | 40 | 56.3 | 370 | 20 | AAW94272 | A mechanotically sen |
| 152 | 45 | 63.4 | 11  | 20 | AAW92679 | Human tachykinin a  | 225 | 40 | 56.3 | 371 | 21 | AAW94272 | Arabidopsis thalia   |
| 153 | 45 | 63.4 | 11  | 22 | AAW92679 | Tachykinins peptid  | 226 | 40 | 56.3 | 411 | 20 | AAW94272 | Human potassium ch   |
| 154 | 45 | 63.4 | 11  | 22 | AAW92679 | Tachykinins peptid  | 227 | 40 | 56.3 | 411 | 20 | AAW94272 | h-TREK1 polypeptid   |
| 155 | 45 | 63.4 | 455 | 20 | AAW92679 | Amino acid sequenc  | 228 | 40 | 56.3 | 411 | 20 | AAW94272 | Mouse h-TREK1 poly   |
| 156 | 44 | 62.0 | 11  | 21 | AAW92679 | Spantide II peptid  | 229 | 40 | 56.3 | 411 | 22 | AAW94272 | Human TREK1          |
| 157 | 43 | 60.6 | 7   | 4  | AAP30468 | Sequence of polype  | 230 | 40 | 56.3 | 429 | 22 | AAW94272 | Human novel protei   |



|     |    |      |      |    |           |                     |     |      |      |      |    |           |                    |
|-----|----|------|------|----|-----------|---------------------|-----|------|------|------|----|-----------|--------------------|
| 231 | 40 | 56.3 | 476  | 22 | AAM23959  | Human EST encoded   | 304 | 38   | 53.5 | 405  | 12 | AAR14404  | Protein 7.2 (1.3-f |
| 232 | 40 | 56.3 | 620  | 21 | AAG35791  | Arabidopsis thalia  | 305 | 38   | 53.5 | 405  | 13 | AAR28840  | HeLa cell fucosylt |
| 233 | 40 | 56.3 | 639  | 22 | AAB95440  | Human protein sequ  | 306 | 38   | 53.5 | 405  | 15 | AAR45937  | A glycosyltransfer |
| 234 | 40 | 56.3 | 725  | 21 | AAG35790  | Arabidopsis thalia  | 307 | 38   | 53.5 | 405  | 18 | AAW13641  | Human alpha(1,3)-f |
| 235 | 40 | 56.3 | 797  | 21 | AAG35789  | Arabidopsis thalia  | 308 | 38   | 53.5 | 405  | 18 | AAW11821  | Human myeloid deri |
| 236 | 40 | 56.3 | 1333 | 15 | AAR63224  | Cobra partial CVF2  | 309 | 38   | 53.5 | 412  | 19 | AAW98765  | H. pylori GHP0 111 |
| 237 | 40 | 56.3 | 1333 | 20 | AAV23730  | Partial cobra veno  | 310 | 38   | 53.5 | 445  | 19 | AAW24153  | Bovine LOX-1 extra |
| 238 | 39 | 54.9 | 7    | 5  | AAP40480  | Substance P analog  | 311 | 38   | 53.5 | 496  | 15 | AAR45938  | A glycosyltransfer |
| 239 | 39 | 54.9 | 12   | 21 | AAV76923  | HIV-1 gp41 inhibit  | 312 | 38   | 53.5 | 530  | 12 | AAR14405  | Protein 1 (1.3-fuc |
| 240 | 39 | 54.9 | 12   | 21 | AAV76931  | HIV-1 gp41 inhibit  | 313 | 38   | 53.5 | 789  | 19 | AAW39927  | Human Arnt recepto |
| 241 | 39 | 54.9 | 16   | 21 | AAV76943  | HIV-1 gp41 inhibit  | 314 | 38   | 53.5 | 901  | 13 | AAAR26790 | Viral enhancing fa |
| 242 | 39 | 54.9 | 18   | 21 | AAV76946  | HIV-1 gp41 inhibit  | 315 | 38   | 53.5 | 901  | 15 | AAAR3963  | VEF. Trichoplusia  |
| 243 | 39 | 54.9 | 20   | 21 | AAV76947  | HIV-1 gp41 inhibit  | 316 | 38   | 53.5 | 1184 | 20 | AAW74445  | Human nucleotide p |
| 244 | 39 | 54.9 | 98   | 22 | AAG76780  | Human colon cancer  | 317 | 38   | 53.5 | 1184 | 21 | AAV66657  | Membrane-bound pro |
| 245 | 39 | 54.9 | 350  | 18 | AAW14532  | Human chimeric fuc  | 318 | 38   | 53.5 | 1184 | 22 | AAU12377  | Human PRO1188 poly |
| 246 | 39 | 54.9 | 359  | 18 | AAW14524  | Human chimeric fuc  | 319 | 38   | 53.5 | 1184 | 22 | AAW65180  | Human PRO1188 (UNQ |
| 247 | 39 | 54.9 | 359  | 18 | AAW14529  | Human chimeric fuc  | 320 | 38   | 53.5 | 1528 | 17 | AAAR95333 | Manduca sexta Bac1 |
| 248 | 39 | 54.9 | 359  | 18 | AAW14531  | Human chimeric fuc  | 321 | 38   | 53.5 | 1528 | 20 | AAW90182  | Manduca sexta Bt t |
| 249 | 39 | 54.9 | 360  | 18 | AAW14512  | Human chimeric fuc  | 322 | 38   | 53.5 | 1721 | 19 | AAW48299  | Cryptosporidium pa |
| 250 | 39 | 54.9 | 360  | 18 | AAW14526  | Human chimeric fuc  | 323 | 38   | 53.5 | 1721 | 21 | AAW11727  | Portion of Cryptos |
| 251 | 39 | 54.9 | 360  | 18 | AAW14513  | Human chimeric fuc  | 324 | 38   | 53.5 | 1837 | 21 | AAW11726  | Cryptosporidium pa |
| 252 | 39 | 54.9 | 360  | 18 | AAW14516  | Human chimeric fuc  | 325 | 37.5 | 52.8 | 220  | 21 | AAW73464  | Human secreted pro |
| 253 | 39 | 54.9 | 361  | 12 | AAR13749  | GDP-Fuc:(beta-D-Ga  | 326 | 37.5 | 52.8 | 487  | 22 | AAW73515  | Human transferase  |
| 254 | 39 | 54.9 | 361  | 15 | AAW45934  | A glycosyltransfer  | 327 | 37.5 | 52.8 | 543  | 19 | AAW72196  | HSV-2 strain SB5 C |
| 255 | 39 | 54.9 | 361  | 18 | AAW23806  | Human alpha 1.3/4   | 328 | 37.5 | 52.8 | 1196 | 19 | AAW72105  | HSV-2 strain SB5 C |
| 256 | 39 | 54.9 | 361  | 18 | AAW13638  | Human alpha(1.3/1,  | 329 | 37   | 52.1 | 5    | 20 | AAW99687  | Substance P analog |
| 257 | 39 | 54.9 | 361  | 18 | AAW14517  | Human chimeric fuc  | 330 | 37   | 52.1 | 6    | 5  | AAW40521  | Sequence of substa |
| 258 | 39 | 54.9 | 361  | 18 | AAW14518  | Human chimeric fuc  | 331 | 37   | 52.1 | 9    | 20 | AAW28521  | Beta-1 integrin ce |
| 259 | 39 | 54.9 | 361  | 18 | AAW14520  | Human chimeric fuc  | 332 | 37   | 52.1 | 9    | 21 | AAW19063  | Anino acid sequenc |
| 260 | 39 | 54.9 | 361  | 22 | AAG64452  | Human Lewis enzyme  | 333 | 37   | 52.1 | 11   | 13 | AAAR21969 | Cyclic substance P |
| 261 | 39 | 54.9 | 374  | 15 | AAAR45939 | A glycosyltransfer  | 334 | 37   | 52.1 | 11   | 13 | AAAR21961 | Cyclic substance P |
| 262 | 39 | 54.9 | 374  | 18 | AAW13642  | Human alpha(1.3)-f  | 335 | 37   | 52.1 | 11   | 20 | AAW92684  | Human tachykinin a |
| 263 | 39 | 54.9 | 463  | 20 | AAV34697  | Chlamydia pneumoni  | 336 | 37   | 52.1 | 11   | 20 | AAW92686  | Human tachykinin a |
| 264 | 39 | 54.9 | 626  | 20 | AAV68292  | Human transcrip tio | 337 | 37   | 52.1 | 11   | 22 | AAW92688  | Tachykinins peptid |
| 265 | 39 | 54.9 | 816  | 19 | AAW68094  | Mouse neuronal PAS  | 338 | 37   | 52.1 | 11   | 22 | AAW91438  | Galanin(1-12)-Pro- |
| 266 | 39 | 54.9 | 824  | 19 | AAW68093  | Human neuronal PAS  | 339 | 37   | 52.1 | 130  | 19 | AAW80717  | S. pneumoniae prot |
| 267 | 39 | 54.9 | 846  | 19 | AAW79533  | Human CLOCK protei  | 340 | 37   | 52.1 | 181  | 21 | AAW15532  | Arabidopsis thalia |
| 268 | 39 | 54.9 | 846  | 20 | AAW84565  | Human HSCLOCK poly  | 341 | 37   | 52.1 | 196  | 21 | AAW44853  | Streptococcus pneu |
| 269 | 39 | 54.9 | 855  | 19 | AAW79529  | Mouse CLOCK protei  | 342 | 37   | 52.1 | 248  | 21 | AAW78087  | Human secreted pro |
| 270 | 39 | 54.9 | 855  | 21 | AAV32214  | Mouse CLOCK protei  | 343 | 37   | 52.1 | 248  | 21 | AAW23892  | Human EST encoded  |
| 271 | 39 | 54.9 | 855  | 21 | AAV32214  | Pinus radiata cell  | 344 | 37   | 52.1 | 248  | 22 | AAE06064  | Human gene 24 enco |
| 272 | 39 | 54.9 | 998  | 1  | AAW25556  | Sequence of neurop  | 345 | 37   | 52.1 | 265  | 20 | AAW02283  | Secreted protein c |
| 273 | 38 | 53.5 | 6    | 9  | AAP80319  | Substance P antago  | 346 | 37   | 52.1 | 272  | 19 | AAW40215  | Human macrophage a |
| 274 | 38 | 53.5 | 6    | 12 | AAR15360  | Substance P analog  | 347 | 37   | 52.1 | 342  | 15 | AAAR63215 | Human alpha-1,3-fu |
| 275 | 38 | 53.5 | 6    | 19 | AAW50977  | Substance P analog  | 348 | 37   | 52.1 | 342  | 15 | AAAR63215 | Human alpha-1,3-fu |
| 276 | 38 | 53.5 | 6    | 19 | AAW50971  | Tachykinins peptid  | 349 | 37   | 52.1 | 359  | 21 | AAW80995  | Murine alpha-1,3-f |
| 277 | 38 | 53.5 | 6    | 22 | AAW91403  | Tachykinins peptid  | 350 | 37   | 52.1 | 359  | 21 | AAW80995  | Human alpha-1,3-fu |
| 278 | 38 | 53.5 | 6    | 22 | AAW91406  | Tachykinins peptid  | 351 | 37   | 52.1 | 388  | 22 | AAW75230  | Drosophila gustato |
| 279 | 38 | 53.5 | 7    | 22 | AAW91400  | Neurokinin recepto  | 352 | 37   | 52.1 | 530  | 21 | AAW43300  | Human ORFX ORF3064 |
| 280 | 38 | 53.5 | 11   | 18 | AAV22688  | Peptide NF1, a su   | 353 | 37   | 52.1 | 565  | 21 | AAW18569  | Amino acid sequenc |
| 281 | 38 | 53.5 | 11   | 19 | AAW60208  | Carboxyfluorescein  | 354 | 37   | 52.1 | 616  | 21 | AAW56941  | Human prostate can |
| 282 | 38 | 53.5 | 11   | 21 | AAV67965  | Peptide #3687 enco  | 355 | 37   | 52.1 | 697  | 21 | AAW27790  | Arabidopsis thalia |
| 283 | 38 | 53.5 | 64   | 22 | AAW17253  | Peptide #3785 enco  | 356 | 37   | 52.1 | 717  | 21 | AAW18609  | Arabidopsis thalia |
| 284 | 38 | 53.5 | 64   | 22 | AAW29748  | Peptide #3626 enco  | 357 | 37   | 52.1 | 735  | 21 | AAW53129  | Arabidopsis thalia |
| 285 | 38 | 53.5 | 64   | 22 | AAW04944  | C. parvum Iowa iso  | 358 | 37   | 52.1 | 735  | 21 | AAW27789  | Arabidopsis thalia |
| 286 | 38 | 53.5 | 91   | 21 | AAW11745  | C. parvum Iowa iso  | 359 | 37   | 52.1 | 745  | 21 | AAW27788  | Arabidopsis thalia |
| 287 | 38 | 53.5 | 124  | 21 | AAW11742  | C. parvum Iowa iso  | 360 | 37   | 52.1 | 755  | 21 | AAW18608  | Arabidopsis thalia |
| 288 | 38 | 53.5 | 128  | 21 | AAW11738  | C. parvum Iowa iso  | 361 | 37   | 52.1 | 755  | 21 | AAW53128  | Arabidopsis thalia |
| 289 | 38 | 53.5 | 130  | 21 | AAW11739  | C. parvum Iowa iso  | 362 | 37   | 52.1 | 765  | 21 | AAW18607  | Arabidopsis thalia |
| 290 | 38 | 53.5 | 130  | 21 | AAW11740  | C. parvum Iowa iso  | 363 | 37   | 52.1 | 765  | 21 | AAW53127  | Arabidopsis thalia |
| 291 | 38 | 53.5 | 138  | 21 | AAW11741  | C. parvum Iowa iso  | 364 | 37   | 52.1 | 874  | 22 | AAW53127  | Arabidopsis thalia |
| 292 | 38 | 53.5 | 150  | 21 | AAW11744  | C. parvum Iowa iso  | 365 | 37   | 52.1 | 958  | 21 | AAW21255  | Murine ADAMTS-9 am |
| 293 | 38 | 53.5 | 159  | 21 | AAW11730  | Cryptosporidium pa  | 366 | 37   | 52.1 | 1073 | 21 | AAW21264  | Human metalloprote |
| 294 | 38 | 53.5 | 159  | 21 | AAW11735  | Cryptosporidium pa  | 367 | 37   | 52.1 | 1081 | 22 | AAW95514  | Human protein sequ |
| 295 | 38 | 53.5 | 175  | 21 | AAW11743  | C. parvum Iowa iso  | 368 | 37   | 52.1 | 1241 | 21 | AAW42626  | Human ORFX ORF2390 |
| 296 | 38 | 53.5 | 249  | 21 | AAW11746  | C. parvum NINC iso  | 369 | 37   | 52.1 | 1306 | 21 | AAW41456  | Human ORFX ORF1220 |
| 297 | 38 | 53.5 | 270  | 17 | AAW99586  | Low density lipopr  | 370 | 37   | 52.1 | 1882 | 22 | AAW72286  | Human ADAMTS-9 am  |
| 298 | 38 | 53.5 | 270  | 20 | AAV24152  | Bovine LDL recepto  | 371 | 37   | 52.1 | 1934 | 22 | AAW72301  | Human ADAMTS-9 alt |
| 299 | 38 | 53.5 | 273  | 17 | AAW99587  | Low density lipopr  | 372 | 37   | 52.1 | 3025 | 22 | AAW99684  | HIV-1 subtype C pr |
| 300 | 38 | 53.5 | 273  | 17 | AAW99588  | Low density lipopr  | 373 | 37   | 52.1 | 10   | 20 | AAW99684  | Substance P analog |
| 301 | 38 | 53.5 | 273  | 20 | AAV24151  | Human LDL recepto   | 374 | 36.5 | 51.4 | 10   | 20 | AAW99684  | Tachykinin peptide |
| 302 | 38 | 53.5 | 405  | 11 | AAW08119  | CDX, a MILA involy  | 375 | 36.5 | 51.4 | 10   | 22 | AAW66675  | Human colon cancer |
| 303 | 38 | 53.5 | 405  | 12 | AAW13752  | GDP-Fuc:beta-D-Gal  | 376 | 36.5 | 51.4 | 147  | 22 | AAW75739  |                    |

377 36.5 51.4 332 21 AAY71125 Human mitogenic re  
378 36.5 51.4 564 9 AAY71119 Human mitogenic re  
379 36 50.7 9 6 AAP50634 Substance P-like p  
380 36 50.7 9 20 AAW92714 Human tachykinin a  
381 36 50.7 9 22 AAB91435 Tachykinins peptid  
382 36 50.7 9 22 AAB91446 Tachykinins peptid  
383 36 50.7 11 22 AAB91437 Tachykinins peptid  
384 36 50.7 12 19 AAW50967 Substance P analog  
385 36 50.7 12 13 AAR29593 HIV-1 gp41 inhibit  
386 36 50.7 13 13 AAR29593 Vertebrate Stromel  
387 36 50.7 16 21 AAY76932 HIV-1 gp41 inhibit  
388 36 50.7 16 21 AAY76943 HIV-1 gp41 inhibit  
389 36 50.7 20 21 AAY76944 HIV-1 gp41 inhibit  
390 36 50.7 39 22 AAM21536 Peptide #11842 enc  
391 36 50.7 39 22 AAM37805 Arabidopsis thalia  
392 36 50.7 46 21 AAG04268 Human secreted pro  
393 36 50.7 51 21 AAB44870 Arabidopsis thalia  
394 36 50.7 51 21 AAG04267 Arabidopsis thalia  
395 36 50.7 83 21 AAG02778 Human secreted pro  
396 36 50.7 89 21 AAG17699 Arabidopsis thalia  
397 36 50.7 108 21 AAG04030 Human secreted pro  
398 36 50.7 110 21 AAG01279 Human secreted pro  
399 36 50.7 163 21 AAY54324 Amino acid sequenc  
400 36 50.7 183 21 AAB08448 A human prostate s  
401 36 50.7 200 21 AAB25378 Pinus radiata cell  
402 36 50.7 218 21 AAB08447 A human prostate s  
403 36 50.7 218 21 AAB08447 A human prostate s  
404 36 50.7 318 22 AAB88540 Haemophilus influe  
405 36 50.7 318 22 AAB60654 Haemophilus influe  
406 36 50.7 359 20 AAY04133 Human Wnt-5b prote  
407 36 50.7 359 21 AAY57271 Wnt-4AF and Wnt-5c  
408 36 50.7 363 20 AAY04134 Human Wnt-5b prote  
409 36 50.7 365 21 AAY70739 Human Wnt-5a prote  
410 36 50.7 365 21 AAY57600 Human Wnt-5a prote  
411 36 50.7 365 22 AAB73619 Wnt-5a tumour supp  
412 36 50.7 372 20 AAY04132 Human Wnt-5b prote  
413 36 50.7 375 21 AAB08449 A human prostate s  
414 36 50.7 388 22 AAG91530 C glutamicum prote  
415 36 50.7 393 21 AAY94425 Human h-TRAAK poly  
416 36 50.7 393 21 AAY94426 Human h-TRAAK poly  
417 36 50.7 398 20 AAY30647 A mechanically sen  
418 36 50.7 414 21 AAB54295 Human pancreatic c  
419 36 50.7 424 22 AAG90605 C glutamicum prote  
420 36 50.7 472 22 AAM00914 Human bone marrow  
421 36 50.7 534 22 AAB98644 Human UDP-glucose:  
422 36 50.7 550 22 AAB80066 Corynebacterium gl  
423 36 50.7 572 22 AAB03506 Human protein kina  
424 36 50.7 611 20 AAW73924 Nucleobase permeas  
425 36 50.7 661 22 AAG89960 C glutamicum prote  
426 36 50.7 729 22 AAM00020 Human Plectin prote  
427 36 50.7 817 22 AAB95492 Human protein sequ  
428 36 50.7 1045 22 AAM00933 Human bone marrow  
429 36 50.7 1493 22 AAB72444 UGCT. Caenorhabdi  
430 36 50.7 1527 22 AAB72436 Rat UGCT. Rattus  
431 36 50.7 2135 22 AAM00019 Human Plectin prote  
432 36 50.7 2594 22 AAM00984 Human bone marrow  
433 35.5 50.0 80 21 AAB51983 Gene 19 human secr  
434 35.5 50.0 80 21 AAB51984 Human secreted pro

## ALIGNMENTS

RESULT 1  
AAP30142  
ID AAP30142 standard; Protein; 11 AA.  
AC AAP30142;  
XX

14-JUN-1992 (first entry)

Sequence of peptides with substance P inhibiting activity.

XX

KW Substance P antagonist; pain therapy; hypertension.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 2  
FT /label= D-P  
FT Modified-site 7  
FT /label= D-W  
FT Misc-difference 8  
FT /label= F,I  
FT Modified-site 9  
FT /label= D-W  
FT Modified-site 11  
FT /label= M,I  
FT /note= "bonded to NH2"  
XX W08301251-A.  
XX 14-APR-1983.  
XX 09-OCT-1981; 81WO-DE00171.  
XX 09-OCT-1981; 81WO-DE00171.  
XX 09-OCT-1981; 81EP-0902802.  
XX (FERR ) FERRING ARZNEIMITTE.  
XX (HORI/) HORIG J.  
XX Horig J;  
XX WPI; 1983-39155K/16 (39155K).  
XX Undeca:peptide derivs. with substance P inhibiting activity -  
XX useful for treating pain and hypertension  
XX Claim 2; Page 18; 25pp; German.  
XX The peptides of the invention are powerful antagonists of Substance  
XX P and so are useful in human and veterinary medicine, for treating  
XX pain and hypertension (esp.) chronic conditions. A 10 microm concn.  
XX of the peptide produced about 50% inhibition at a Substance P concn. of  
XX 7.5-20 nanom.  
XX Sequence 11 AA;  
SQ  
Query Match 100.0%; Score 71; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred No. 8.5e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQQQFWLM 11  
Db 1 rpkpqqqfwlm 11  
RESULT 2  
AAP80317  
ID AAP80317 standard; protein; 11 AA.  
XX  
AC AAP80317;  
XX  
DT 14-SEP-1990 (first entry)  
XX  
DE Sequence of neuropeptide antagonist E which binds with polypeptide  
DE receptor for bombesin type polypeptides.  
XX  
KW Spantide; neuropeptide; polypeptide receptor; cancer diagnosis;  
KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides;  
KW antagonist E.  
XX  
XX Swiss 3T3 cells.  
XX OS  
XX Key Location/Qualifiers  
FH Misc-difference 2  
FT

FT /label=OTHER  
 FT /note="DPro"  
 FT  
 FT Misc-difference 7  
 FT /label=OTHER  
 FT /note="DTrp"  
 FT  
 FT Misc-difference 9  
 FT /label=OTHER  
 FT /note="DTrp"  
 FT  
 FT Misc-difference 11  
 FT /label=OTHER  
 FT /note="Met-NH2"  
 FT  
 XX W08807551-A.  
 XX  
 XX  
 PD 06-OCT-1988.  
 XX  
 XX 31-MAR-1988; 88WO-GB00255.  
 PF  
 XX 25-NOV-1987; 87GB-0027638.  
 XX  
 XX (IMCR ) IMPERIAL CANCER RES.  
 PA  
 XX Rosengurt E, Zachary I, Woll P;  
 PI  
 XX WPI; 1988-292842/41.  
 DR  
 XX  
 XX  
 XX  
 PT New polypeptide receptor for bombesin type polypeptide(s) -  
 PT is isolated from surface of Swiss 3T3 cells, and antibodies and  
 PT antagonists are useful for treating uncontrolled cell proliferation  
 PT  
 XX  
 PS Disclosure; Table 2; 42pp; English.  
 XX

CC The patent claims a polypeptide isolated from the surface of Swiss 3T3  
 CC cells which binds selectively with polypeptides of the bombesin type and  
 CC binds with antagonist A and antagonist D. Antagonist A is a  
 CC commercially available structural variant of substance P, known as  
 CC [D-ArG1, D-Pro2, D-Trp7,9, Leu11] substance P. It is also known as  
 CC [D-Pro2] spantide. Antagonist B is also commercially available structural  
 CC variant of substance P, known as [D-Phe5] spantide. Substance P is an  
 CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten  
 CC substance P antagonists (see AAP0313-80322) were tested for their  
 CC ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue  
 CC of bombesin in Swiss 3T3 cells). Antagonist D was clearly the most  
 CC potent GRP antagonist. Peptides B, C, D, E, F, G, H, J and K were less  
 CC potent than either A or D. Spantide (B) had no antagonist activity even  
 CC at 100 uM. Polypeptide antagonists A and B and novel variants are useful  
 CC for diagnosis and therapy, esp. of cancers where uncontrolled cell  
 CC growth is associated with disorders of proteins of the bombesin family.  
 XX

SQ Sequence 11 AA;

Query Match 100.0%; Score 71; DB 9; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQOWFWLM 11  
 |||||  
 Db 1 rpkpqowfwlm 11

RESULT 3

ID AAW50969  
 XX AAW50969 standard; peptide; 11 AA.  
 AC AAW50969;

XX  
 XX 31-JUL-1998 (first entry)  
 DT  
 XX

DE Substance P analogue, [D-Pro2,D-Trp7,9].

XX Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
 KW Substance P; cancer; inhibition; growth hormone releasing factor;  
 KW

KW spantide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2 /note= "D-form residue"  
 FT  
 FT Misc-difference 7 /note= "D-form residue"  
 FT  
 FT Misc-difference 9 /note= "D-form residue"  
 FT  
 FT Modified-site 11 /note= "D-form residue"  
 FT  
 FT /note= "C-terminal amide"  
 XX  
 XX EP835662-A2.  
 PN  
 XX  
 PD 15-APR-1998.  
 XX  
 XX 11-DEC-1996; 96EP-0309012.  
 PF  
 XX 08-OCT-1996; 96US-0727679.  
 PR  
 PR 16-AUG-1996; 96IN-0001822.  
 XX  
 XX (NAIM-) NAT INST IMMUNOLOGY.  
 PA  
 XX Jaggi M, Mukherjee R;  
 PI  
 XX WPI; 1998-208959/19.  
 DR  
 XX

XX Composition containing analogues of vasoactive intestinal peptide,  
 PT somatostatin - bombesin and substance P, for treatment of tumours  
 PT and for inhibiting over-expression of these peptide(s)  
 PT  
 XX Disclosure; Page 13; 49pp; English.  
 PS  
 XX The invention relates to a new composition which comprises: (i) the  
 CC somatostatin analogue SOM2 AGCKNFFdWKPTSDC (3-14 disulphide bridge),  
 CC and (ii) at least 4 of the peptides: antagonist of vasoactive  
 CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP  
 CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin  
 CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are  
 CC more general compositions containing peptide analogues of somatostatin,  
 CC VIP, bombesin and substance P. The compositions are used in human or  
 CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour  
 CC or cancer cells, particularly for treatment of leukaemia, lymphoma,  
 CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,  
 CC breast, kidney or particularly rectum and colon, and (b) to prevent,  
 CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
 CC cells express receptors for VIP, somatostatin, bombesin and/or substance  
 CC P. The present sequence represents a substance P analogue.  
 XX

SQ Sequence 11 AA;

Query Match 100.0%; Score 71; DB 19; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 8.5e-05;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQOWFWLM 11  
 |||||  
 Db 1 rpkpqowfwlm 11

RESULT 4

AAW92656  
 ID AAW92656 standard; peptide; 11 AA.  
 XX  
 AC AAW92656;

XX 30-APR-1999 (first entry)  
 DT  
 XX

DE Human tachykinin agonist beta-amyloid peptide fragment #2.  
 XX

/note= "D-form residue"  
 Misc-difference 7  
 /note= "D-form residue"  
 Misc-difference 9  
 /note= "D-form residue"  
 Modified-site 11  
 /label= OTHER  
 /note= "C-terminal amide"  
 WO200130371-A2.  
 03-MAY-2001.  
 27-OCT-2000; 2000WO-US29789.  
 28-OCT-1999; 99US-0428692.  
 (NEWWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
 Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
 WPI; 2001-397593/42.

27-OCT-2000; 2000WO-US29789.  
28-OCT-1999; 99US-0428692.  
(NEWE-) NEW ENGLAND MEDICAL CENT HOSPITAL  
Carr DB, Lipkowski AW, Kream R, Misick  
WPI; 2001-397593/42.

invention.                      11 AA: Sequence

```

Very Match          100.0%; Score 71;
1st Local Similarity 100.0%; Pred. No. 8
Archives 11: Conservative 0; Mismatches

```

```

1 RPKPQQWFWM 11
  |||||
1 rpkpgawfwlm 11

```

LT 6  
1413  
AAB91413 standard: peptide: 11 AA.

AAB91413;  
22-JUN-2001 (first entry)

Tachykinins peptide SEQ ID NO:589.

Homo sapiens.  
Synthetic.

23-NOV-2000.  
17-MAY-2000. 2000WO-11S13576

17 MAY 2000, 2000WG0313370.  
17-MAY-1999; 99US-0134406.  
10-SEP-1999; 99US-0153406.  
15-OCT-1999; 99US-0159783.

|                                                                     |                         |
|---------------------------------------------------------------------|-------------------------|
| Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment |                         |
| Alzheimer's disease; Down's syndrome; amyloidosis; human;           |                         |
| hereditary cerebral haemorrhage; non-inherited congenital angiop    |                         |
| Homo sapiens.                                                       |                         |
| Key                                                                 | Location/Qualifiers     |
| Misc-difference 2                                                   | /note= "D-form residue" |
| Misc-difference 7                                                   | /note= "D-form residue" |
| Misc-difference 9                                                   | /note= "D-form residue" |
| US5876948-A.                                                        |                         |
| 02-MAR-1999.                                                        |                         |
| 27-JUL-1991;                                                        | 91US-0737371.           |
| 29-JUL-1991;                                                        | 91US-0737371.           |
| 27-JUL-1990;                                                        | 90US-0539173.           |

US5876948-A.  
02-MAR-1999.  
27-JUL-1991; 91US-0737371.  
29-JUL-1991; 91US-0737371.  
27-JUL-1990; 90US-0559173.

This invention describes a method for s a neurotoxin. The method involves incub

neuronal cells and a beta-amyloid peptide used for identifying compounds for treatment of Alzheimer's disease. The synthesis of undesirable build up of beta-amyloid protein in Down's syndrome, and the syndromes of haemophilia, haemorrhage, and non-inherited congenital haemorrhage. AAN92655-W92731 are tachykinin-like peptides. AAN92655-W92731 are tachykinin-like peptides. AAN92655-W92731 are tachykinin-like peptides.

Sequence 11 AA;

|                         |        |           |
|-------------------------|--------|-----------|
| every Match             | 100.0% | Score 71; |
| tt Local Similarity     | 100.0% | Pred. No. |
| atches 11: Conservative | 0:     | Mismatch  |

1 RPKPQQWFWM 11  
 |||||  
 1 rpkpqaawfwlm 11

T 5  
881  
AAB98801 standard; Peptide; 11 AA.

AAB98881;  
14-AUG-2001 (first entry)  
Chimeric analgesic peptide #17

Opioïd receptor binding; nociceptive re-  
pain; chimeric peptide.

Synthetic.

| Key               | Location/Qualifiers |
|-------------------|---------------------|
| Misc-difference 2 |                     |

XX PA (CONJ-) CONJUCHEM INC.  
XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX DR WPI; 2001-112059/12.  
XX PT Modifying and attaching therapeutic peptides to albumin prevents  
XX PT peptidase degradation, useful for increasing length of in vivo activity  
XX PS Disclosure; Page 392; 733pp; English.  
XX CC The present invention describes a modified therapeutic peptide (I)  
XX CC comprising a therapeutically active amino acid region (III) and a  
XX CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
XX CC a less therapeutically active amino acid region (IV), which covalently  
XX CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX CC factors and neurotransmitters, to protect them from peptidase activity  
XX CC in vivo for the treatment of various disorders. Endogenous therapeutic  
XX CC peptides are not suitable as drug candidates as they require frequent  
XX CC administration due to rapid degradation by peptidases in the body.  
XX CC Modifying and attaching therapeutic peptides to albumin prevents or  
XX CC reduces the action of peptidases to increase length of activity (half  
XX CC life) and specificity as bonding to large molecules decreases  
XX CC intracellular uptake and interference with physiological processes.  
XX CC AAB90829 to AAB92441 represent peptides which can be used in the  
XX CC exemplification of the present invention.  
XX SQ Sequence 11 AA;  
  
Query Match 100.0%; Score 71; DB 22; Length 11;  
Best Local Similarity 100.0%; Pred. No. 8.5e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 RPKPQQWFWM 11  
Db 1 rpkpqwfwm 11  
  
RESULT 7  
AAB98882  
ID AAB98882 standard; Peptide; 12 AA.  
XX AC AAB98882;  
XX DT 14-AUG-2001 (first entry)  
XX DE Chimeric analgesic peptide #38.  
XX KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
XX KW pain; chimeric peptide.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "D-form residue"  
FT FT Misc-difference 7 /note= "D-form residue"  
FT FT Misc-difference 9 /note= "D-form residue"  
FT FT Modified-site 12 /label= OTHER  
FT FT /note= "C-terminal amide"  
XX PN W0200130371-A2.  
XX PD 03-MAY-2001.  
XX XX

PF 27-OCT-2000; 2000WO-US29789.  
XX PR 28-OCT-1999; 99US-0428692.  
XX PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX DR WPI; 2001-397593/42.  
XX PT New chimeric peptides used for treating pain comprise opioid receptor  
XX PT binding group and nociceptive receptor binding group  
XX PS Claim 10; Page 16; 34pp; English.  
XX CC The present invention describes a number of chimeric peptides comprising  
XX CC an opioid receptor binding moiety and a nociceptive receptor binding  
XX CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
XX CC opioid receptor based peptides alone, tolerance does not result from  
XX CC their long-term use. The present sequence is one of the peptides of the  
XX CC invention.  
XX SQ Sequence 12 AA;  
  
Query Match 100.0%; Score 71; DB 22; Length 12;  
Best Local Similarity 100.0%; Pred. No. 9.2e-05;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 RPKPQQWFWM 11  
Db 1 rpkpqwfwm 11  
  
RESULT 8  
AAP40479  
ID AAP40479 standard; peptide; 11 AA.  
XX AC AAP40479;  
XX DT 27-NOV-1991 (first entry)  
XX DE Substance P analogue.  
XX KW Substance P; analogue; antiinflammatory agent; analgesic.  
XX PN US4481139-A.  
XX PD 06-NOV-1984.  
XX PF 13-APR-1983; 83US-0484646.  
XX PR 13-APR-1983; 83US-0484646.  
XX PA (UYTE-) UNIVERSITY OF TEXAS SYSTEM.  
XX PI Folkers K, Ji-cheng X;  
XX DR WPI; 1984-294258/47.  
XX PT Peptide analogues of substance P - useful as antagonists, e.g. as  
XX PT antiinflammatory agents and analgesics.  
XX PS Claim 1; page 5; 5pp; English.  
XX CC The peptide is a D-Arg1, D-Trp7, D-Trp9, Leu11 analogue of substance  
XX CC P. The peptide is a substance P antagonist with higher activity than  
XX CC known substance P analogues. It may be used as a biological  
XX CC research tool, ophthalmological antiinflammatory agent and analgesic.  
XX SQ Sequence 11 AA;

Query Match 95.8%; Score 68; DB 5; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00024;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
| | | | | | | | | |  
Db 1 rpkpqwfll 11

RESULT 9

AAP80313  
ID AAP80313 standard; protein; 11 AA.  
XX  
AC AAP80313;  
XX  
DT 14-SEP-1990 (first entry)  
XX  
DE Sequence of neuropeptide antagonist A which binds with polypeptide  
DE receptor for bombesin type polypeptides.  
XX  
KW Spantide; neuropeptide; polypeptide receptor; bombesin; cancer diagnosis;  
KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides;  
KW antagonist A.  
XX  
OS Swiss 3T3 cells.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1 /label=OTHER  
FT /note="Darg"  
FT Misc-difference 2 /label=OTHER  
FT /note="DPro"  
FT Misc-difference 7 /label=OTHER  
FT /note="DTrp"  
FT Misc-difference 9 /label=OTHER  
FT /note="DTrp"  
FT Misc-difference 11 /label=OTHER  
FT /note="Leu-NH2"

WO8807551-A.  
XX  
XX  
XX 06-OCT-1988.  
XX  
XX 31-MAR-1988; 88WO-GB00255.  
XX  
XX 25-NOV-1987; 87GB-0027638.  
XX  
XX (IMCR ) IMPERIAL CANCER RES.  
XX  
XX Rosengurt E, Zachary I, Woll P;  
XX  
XX WPI; 1988-292842/41.  
XX  
XX New polypeptide receptor for bombesin type polypeptide(s) -  
PT is isolated from surface of Swiss 3T3 cells, and antibodies and  
PT antagonists are useful for treating uncontrolled cell proliferation  
PS Disclosure; Table 2; 42pp; English.  
XX  
XX The patent claims a polypeptide isolated from the surface of Swiss 3T3  
CC cells which binds selectively with polypeptides of the bombesin type and  
CC binds with antagonist A and antagonist D. Antagonist A is a  
CC commercially available structural variant of substance P, known as  
CC [D-Arg1, D-Pro2, D-Trp7,9, Leu11] substance P. It is also known as  
CC [D-Pro2] spantide. Antagonist B is also commercially available structural  
CC variant of substance P, known as [D-Phe5] spantide. Substance P is an  
CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten  
CC substance P antagonists (see AAP80313-80322) were tested for their  
CC ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue

CC of bombesin in Swiss 3T3 cells). Antagonist D was clearly the most  
CC potent GRP antagonist. Peptides B, C, D, E, F, G, H, J and K were less  
CC potent than either A or D. Spantide (B) had no antagonist activity even  
CC at 100 uM. Polypeptide antagonists A and D and novel variants are useful  
CC for diagnosis and therapy, esp. of cancers where uncontrolled cell  
CC growth is associated with disorders of proteins of the bombesin family.  
XX  
SQ Sequence 11 AA;

Query Match 95.8%; Score 68; DB 9; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00024;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
| | | | | | | | | |  
Db 1 rpkpqwfll 11

RESULT 10

AAP80314  
ID AAP80314 standard; protein; 11 AA.  
XX  
AC AAP80314;  
XX  
DT 14-SEP-1990 (first entry)  
XX  
DE Sequence of neuropeptide antagonist B which binds with polypeptide  
DE receptor for bombesin type polypeptides.  
XX  
KW Spantide; neuropeptide; polypeptide receptor; bombesin; cancer diagnosis;  
KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides;  
KW antagonist B.  
XX  
OS Swiss 3T3 cells.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1 /label=OTHER  
FT /note="Darg"  
FT Misc-difference 7 /label=OTHER  
FT /note="DTrp"  
FT Misc-difference 1 /label=OTHER  
FT /note="DTrp"  
FT Misc-difference 14 /label=OTHER  
FT /note="Leu-NH2"

WO8807551-A.  
XX  
XX  
XX 06-OCT-1988.  
XX  
XX 31-MAR-1988; 88WO-GB00255.  
XX  
XX 25-NOV-1987; 87GB-0027638.  
XX  
XX (IMCR ) IMPERIAL CANCER RES.  
XX  
XX Rosengurt E, Zachary I, Woll P;  
XX  
XX WPI; 1988-292842/41.  
XX  
XX New polypeptide receptor for bombesin type polypeptide(s) -  
PT is isolated from surface of Swiss 3T3 cells, and antibodies and  
PT antagonists are useful for treating uncontrolled cell proliferation  
PS Disclosure; Table 2; 42pp; English.  
XX  
XX The patent claims a polypeptide isolated from the surface of Swiss 3T3  
CC cells which binds selectively with polypeptides of the bombesin type and  
CC binds with antagonist A and antagonist D. Antagonist A is a  
CC commercially available structural variant of substance P, known as  
CC [D-Arg1, D-Pro2, D-Trp7,9, Leu11] substance P. It is also known as  
CC [D-Pro2] spantide. Antagonist B is also commercially available structural  
CC variant of substance P, known as [D-Phe5] spantide. Substance P is an  
CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten  
CC substance P antagonists (see AAP80313-80322) were tested for their  
CC ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue

CC commercially available structural variant of substance P, known as  
CC [D-Ar<sup>1</sup>, D-Pro<sup>2</sup>, D-Trp<sup>7,9</sup>, Leu<sup>11</sup>] substance P. It is also known as  
CC [D-Pro<sup>2</sup>] spantide. Antagonist B is also commercially available structural  
CC variant of substance P, known as [D-Phe<sup>5</sup>] spantide. Substance P is an  
CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten  
CC substance P antagonists (see AAR0313-80322) were tested for their  
CC ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue  
CC of bombesin in Swiss 3T3 cells). Antagonist D was clearly the most  
CC potent GRP antagonist. Peptides B, C, D, E, F, G, H, J and K were less  
CC potent than either A or D. Spantide (B) had no antagonist activity even  
CC at 100 nM. Polypeptide antagonists A and D and novel variants are useful  
CC for diagnosis and therapy, esp. of cancers where uncontrolled cell  
CC growth is associated with disorders of proteins of the bombesin family.

XX Sequence 11 AA;

Query Match 95.8%; Score 68; DB 9; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00024;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQOQFWLM 11  
|||||  
Db 1 rpkpqgfwll 11

## RESULT 11

AAR05856  
ID AAR05856 standard; protein; 11 AA.

XX

AC AAR05856;

XX

DT 07-SEP-1990 (first entry)

XX

DE D-arginine 1, D-proline 2, D-tryptophan 7,9, Leucine 11,

DE -substance P angiotensin antagonist.

XX

KW Angiotensin; ectopic hormone; mas oncogene; cancer;

KW Neuroblastoma; neuroendocrine.

XX

OS Synthetic.

XX

FT Key Location/Qualifiers

FT Modified-site 1

FT /label=Dextrorotatory form.

FT Modified-site 2

FT /label=Dextrorotatory form.

FT Modified-site 7

FT /label=Dextrorotatory form.

FT Modified-site 9

FT /label=Dextrorotatory form.

XX

XX WO9003181-A.

PN

XX

PD 05-APR-1990.

XX

PF 22-SEP-1989; 89WO-0001121.

XX

XX 24-SEP-1988; 88GB-0022483.

PR

XX (MEDI-) MED RES COUNCIL.

XX

XX Hanley MR, Goedert M;

PI

XX WPI; 1990-132106/17.

DR

XX Use of substances which block the activity of angiotensin -

PT for treatment or prevention of tumour development or ectopic

PT hormone prodn.

XX

XX Claim 8; Page 19; 23pp; English.

PS

XX Peptide blocks biological activity of angiotensin and is active

CC

CC against the mas oncogene, retarding tumour growth, esp  
CC neuroendocrine and neuroblastoma tumours.

XX Sequence 11 AA;

Query Match 95.8%; Score 68; DB 11; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00024;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQOQFWLM 11  
|||||  
Db 1 rpkpqgfwll 11

## RESULT 12

AAR11144

ID AAR11144 standard; Protein; 11 AA.

XX

AC AAR11144;

XX

DT 21-MAY-1991 (first entry)

XX

DE Substance P analogue.

XX

KW Anti-proliferation agent; neurogenetic inflammation; fibroblasts;

KW agonist.

XX

OS Synthetic.

XX

FT Key Location/Qualifiers

FT Modified-site 1

FT /label= D-Arg

FT Modified-site 7

FT /label= D-Trp

FT Modified-site 9

FT /label= D-Trp

FT Modified-site 9..10

FT /label= non-peptide bond

FT /note= "Trp-L[CH2NH]-Trp"

FT Modified-site 11

FT /label= Nle

XX

XX WO9102745-A.

PN

XX

PD 07-MAR-1991.

XX

PF 16-AUG-1990; 90WO-US04633.

XX

XX 16-AUG-1989; 89US-0394727.

PR

XX (TULA ) TULANE E FUND ADMINISTRA.

PA

XX Coy DH, Moreau JP;

PI

XX WPI; 1991-087240/12.

DR

XX Modified linear peptide analogue of natural substance P - acts as

PT competitive inhibitor of substance P and is used for treating

PT neuro genetic inflammation and as anti-proliferative agent.

XX

XX Claim 11; Page 34; 40pp; English.

PS

XX The peptide has a non-peptide bond introduced between Trp<sup>9</sup> and

CC Leu<sup>10</sup>. This may alternatively be positioned between Leu<sup>10</sup> and

CC Nle<sup>11</sup>. For prepn., a benzhydrylamine resin was coupled to Boc-Leu.

CC Boc-Leu aldehyde was dissolved in 5 ml DMF and added to the resin

CC TFA salt suspension followed by addn. of NaCNBH<sub>3</sub> and stirring for

CC one hour. The remaining amino acids were then coupled successively.

CC In tests the peptide inhibited P-stimulated amylase release from

CC pancreatic acini.

CC See also AAR11143.

XX

SQ Sequence 11 AA;

Query Match 95.8%; Score 68; DB 12; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00024;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
| | | | | | | | | | |  
Db 1 rpkpqgwfll 11

RESULT 13  
AAW50966  
ID AAW50966 standard; peptide; 11 AA.  
XX AC AAW50966;  
XX DT 31-JUL-1998 (first entry)  
XX DE Substance P analogue, spantide I.  
XX KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
KW Substane P; cancer; inhibition; growth hormone releasing factor;  
KW spantide.  
XX OS Synthetic.  
XX XX

Key Location/Qualifiers  
FH Misc-difference 1 /note= "D-form residue"  
FT Misc-difference 7 /note= "D-form residue"  
FT Misc-difference 9 /note= "D-form residue"  
FT Misc-difference 11 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
XX EP835662-A2.  
XX PD 15-APR-1998.  
XX PF 11-DEC-1996; 96EP-0309012.  
XX PR 08-OCT-1996; 96US-0727679.  
XX PR 16-AUG-1996; 96IN-0001822.  
XX PA (NAIM-) NAT INST IMMUNOLOGY.  
XX PI Jaggi M, Mukherjee R;  
XX WI; 1998-208959/19.  
XX XX  
XX Composition containing analogues of vasoactive intestinal peptide,  
PT somatostatin - bombesin and substance P, for treatment of tumours  
PT and for inhibiting over-expression of these peptide(s)  
XX XX  
XX Disclosure; Page 13; 49pp; English.

The invention relates to a new composition which comprises: (1) the  
CC somatostatin analogue SOM2 AGCKNFFQWKPTSDC (3-14 disulphide bridge),  
CC and (ii) at least 4 of the peptides: antagonist of vasoactive  
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP  
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin  
CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are  
CC more general compositions containing peptide analogues of somatostatin,  
CC VIP, bombesin and substance P. The compositions are used in human or  
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour  
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,  
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,  
CC breast, kidney or particularly rectum and colon, and (b) to prevent,  
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
CC cells express receptors for VIP, somatostatin, bombesin and/or substance

CC P. The present sequence represents a substance P analogue, spantide I.  
XX  
SQ Sequence 11 AA;

Query Match 95.8%; Score 68; DB 19; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00024;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
| | | | | | | | | | |  
Db 1 rpkpqgwfll 11

RESULT 14  
AAW50958  
ID AAW50958 standard; peptide; 11 AA.  
XX AC AAW50958;  
XX DT 31-JUL-1998 (first entry)  
XX DE Substance P analogue, [D-Arg1,D-Pro2,D-Trp7,9,Leu11]-Substance P.  
XX KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
KW Substane P; cancer; inhibition; growth hormone releasing factor.  
XX OS Synthetic.  
XX XX

Key Location/Qualifiers  
FH Misc-difference 1 /note= "D-form residue"  
FT Misc-difference 2 /note= "D-form residue"  
FT Misc-difference 7 /note= "D-form residue"  
FT Misc-difference 9 /note= "D-form residue"  
FT Misc-difference 11 /note= "D-form residue"  
FT Modified-site 11 /note= "C-terminal amide"  
XX EP835662-A2.  
XX PD 15-APR-1998.  
XX PF 11-DEC-1996; 96EP-0309012.  
XX PR 08-OCT-1996; 96US-0727679.  
XX PR 16-AUG-1996; 96IN-0001822.  
XX PA (NAIM-) NAT INST IMMUNOLOGY.  
XX PI Jaggi M, Mukherjee R;  
XX WI; 1998-208959/19.  
XX XX  
XX Composition containing analogues of vasoactive intestinal peptide,  
PT somatostatin - bombesin and substance P, for treatment of tumours  
PT and for inhibiting over-expression of these peptide(s)  
XX XX  
XX Disclosure; Page 12; 49pp; English.

The invention relates to a new composition which comprises: (1) the  
CC somatostatin analogue SOM2 AGCKNFFQWKPTSDC (3-14 disulphide bridge),  
CC and (ii) at least 4 of the peptides: antagonist of vasoactive  
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP  
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin  
CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are  
CC more general compositions containing peptide analogues of somatostatin,  
CC VIP, bombesin and substance P. The compositions are used in human or  
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour  
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,  
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,  
CC breast, kidney or particularly rectum and colon, and (b) to prevent,  
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
CC cells express receptors for VIP, somatostatin, bombesin and/or substance



CC breast, kidney or particularly rectum and colon, and (b) to prevent,  
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
CC cells express receptors for VIP, somatostatin, bombesin and/or substance  
CC P. The present sequence represents a substance P analogue.  
XX  
SQ Sequence 11 AA;

Query Match 95.8%; Score 68; DB 19; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00024;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
|||||  
Db 1 rpkpqgfwll 11

## RESULT 15

AAW92657  
ID AAW92657 standard; peptide; 11 AA.

XX AC AAW92657;

XX DT 30-APR-1999 (first entry)

XX DE Human tachykinin agonist beta-amyloid peptide fragment #3.

XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
XX KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
XX KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX OS Homo sapiens.

XX PH Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"  
FT Misc-difference 7

FT Misc-difference 9 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

XX FN US5876948-A.

XX PD 02-MAR-1999.

XX PF 27-JUL-1991; 91US-0737371.

XX PR 29-JUL-1991; 91US-0737371.

XX PR 27-JUL-1990; 90US-0559173.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI Yankner BA;

XX DR WPI; 1999-189630/16.

XX PT Screening for neurotoxin inhibitors - by testing compounds for their  
XX PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells

XX PS Disclosure: Column 11-12; 28pp; English.

XX CC This invention describes a method for screening compounds for inhibiting  
XX CC a neurotoxin. The method involves incubating tachykinin agonists with  
XX CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
XX CC used for identifying compounds for treating diseases characterised by an  
XX CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
XX CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
XX CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
XX CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
XX CC beta-amyloid peptide fragments.

XX SQ Sequence 11 AA;

Query Match 95.8%; Score 68; DB 22; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00024;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.8%; Score 68; DB 20; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00024;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
|||||  
Db 1 rpkpqgfwll 11

## RESULT 16

AAW91434  
ID AAW91434 standard; peptide; 11 AA.

XX AC AAW91434;

XX DT 22-JUN-2001 (first entry)

XX DE Tachykinins peptide SEQ ID NO:610.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
XX KW blood component; modification; succinimidyl; maleimido group; amino;  
XX KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200069900-A2.

XX PD 23-NOV-2000.

XX PF 17-MAY-2000; 2000WO-US13576.

XX PR 17-MAY-1999; 99US-0134406.

XX PR 10-SEP-1999; 99US-0153406.

XX PR 15-OCT-1999; 99US-0159783.

XX PA (CONJ-) CONJUCHEM INC.

XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX DR WPI; 2001-112059/12.

XX PT Modifying and attaching therapeutic peptides to albumin prevents  
XX PT peptidase degradation, useful for increasing length of in vivo activity

XX PS Disclosure; Page 398; 733pp; English.

XX CC The present invention describes a modified therapeutic peptide (I)  
XX CC comprising a therapeutically active amino acid region (III) and a  
XX CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
XX CC a less therapeutically active amino acid region (IV), which covalently  
XX CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX CC factors and neurotransmitters, to protect them from peptidase activity  
XX CC in vivo for the treatment of various disorders. Endogenous therapeutic  
XX CC peptides are not suitable as drug candidates as they require frequent  
XX CC administration due to rapid degradation by peptidases in the body.  
XX CC Modifying and attaching therapeutic peptides to albumin prevents or  
XX CC reduces the action of peptidases to increase length of activity (half  
XX CC life) and specificity as bonding to large molecules decreases  
XX CC intracellular uptake and interference with physiological processes.  
XX CC AAW90829 to AAW92441 represent peptides which can be used in the  
XX CC exemplification of the present invention.

XX SQ Sequence 11 AA;

QY 1 RPKPQQWFWM 11  
Db 1 rpkpqgfwll 11

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 17  
AAR28680  
ID AAR28680 standard; Protein; 24 AA.  
XX AC AAR28680;  
XX DT 22-MAR-1993 (first entry)  
XX DE Galanin(1-12)-Pro-Spantide amide (C7).  
XX KW Receptor; Substance P; insulin; growth hormone;  
KW acetylcholine; dopamine; somatostatin; noradrenaline;  
KW endocrinology; food intake; neurology; psychiatry;  
KW Alzheimer-type senile dementia; schizophrenia;  
KW intestinal diseases.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Misc-difference 14 /note= "D-form residue"  
FT Misc-difference 20 /note= "D-form residue"  
FT Misc-difference 22 /note= "D-form residue"  
FT Peptide 1..12 /label= galanin(1-12)  
FT Peptide 14..24 /label= spantide  
XX EP514361-A.  
XX 19-NOV-1992.  
XX 14-MAY-1992; 92EP-0850108.  
XX 15-MAY-1991; 91SE-0001472.  
XX (ASTR ) ASTRA AB.  
XX Ahren B, Bartfai T, Consolo S, Hoekfelt T, Land T;  
XX Langel U, Lindskog S, Wiesenfeld-Hallin Z;  
XX WPI; 1992-384184/47.  
XX New galanin antagonist peptide(s) - used for treating  
XX Alzheimer's-type senile dementia, schizophrenia, analgesia and  
XX intestinal diseases  
XX Disclosure; Page 7; 21pp; English.  
XX The C-terminal of this peptide is amidated. MW= 2827; IC50= 0.2nM.  
XX The peptides given in AAR28679-90 are used to treat disorders in  
XX mammals caused by the function of galanin at its receptor. The  
XX peptides may be useful in the regulation of insulin release, growth  
XX hormone release, acetylcholine release, dopamine release, substance  
XX P release, somatostatin release and noradrenaline release. They are  
XX useful in endocrinology, food intake, neurology and psychiatry, and  
XX to treat Alzheimer-type senile dementia, schizophrenia, intestinal  
XX diseases, and in analgesia. Dosage is 0.01-1000, pref. 0.1-1000  
XX microg/kg body wt.  
XX Sequence 24 AA;  
XX Query Match 95.8%; Score 68; DB 13; Length 24;  
XX Best Local Similarity 90.9%; Pred. No. 0.00051;

QY 1 RPKPQQWFWM 11  
Db 14 rpkpqgfwll 24

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 18  
AAB92023  
ID AAB92023 standard; Peptide; 24 AA.  
XX AC AAB92023;  
XX DT 22-JUN-2001 (first entry)  
XX DE Galanin peptide SEQ ID NO:1199.  
XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidy; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX OS Homo sapiens.  
XX OS Synthetic.  
XX WO200069900-A2.  
XX 23-NOV-2000.  
XX 17-MAY-2000; 2000WO-US13576.  
XX 17-MAY-1999; 99US-0134406.  
XX 10-SEP-1999; 99US-0153406.  
XX 15-OCT-1999; 99US-0159783.  
XX (CONJ-) CONJUCHEM INC.  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX WPI; 2001-112059/12.  
XX Modifying and attaching therapeutic peptides to albumin prevents  
XX peptidase degradation, useful for increasing length of in vivo activity  
XX Disclosure; Page 586; 733pp; English.  
XX The present invention describes a modified therapeutic peptide (I)  
XX comprising a therapeutically active amino acid region (III) and a  
XX reactive group (II) (e.g. succinimidy and maleimido groups) attached to  
XX a less therapeutically active amino acid region (IV), which covalently  
XX bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX factors and neurotransmitters, to protect them from peptidase activity  
XX in vivo for the treatment of various disorders. Endogenous therapeutic  
XX peptides are not suitable as drug candidates as they require frequent  
XX administration due to rapid degradation by peptidases in the body.  
XX Modifying and attaching therapeutic peptides to albumin prevents or  
XX reduces the action of peptidases to increase length of activity (half  
XX life) and specificity as bonding to large molecules decreases  
XX intracellular uptake and interference with physiological processes.  
XX AAB90829 to AAB92441 represent peptides which can be used in the  
XX exemplification of the present invention.  
XX Sequence 24 AA;  
XX Query Match 95.8%; Score 68; DB 22; Length 24;  
XX Best Local Similarity 90.9%; Pred. No. 0.00051;  
XX Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
Db 1 rpkpqgfwll 11

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.8%; Score 68; DB 13; Length 24;  
Best Local Similarity 90.9%; Pred. No. 0.00051;

Db 14 rpkpqwfwl1 24

## RESULT 19

AAB92031  
ID AAB92031 standard; Peptide; 24 AA.

XX AC AAB92031;

XX DT 22-JUN-2001 (first entry)

XX DE Galanin peptide SEQ ID NO:1207.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
XX KW blood component; modification; succinimidyl; maleimido group; amino;  
XX KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200069900-A2.

XX PD 23-NOV-2000.

XX PF 17-MAY-2000; 2000WO-US13576.

XX PR 17-MAY-1999; 99US-0134406.

XX PR 10-SEP-1999; 99US-0153406.

XX PR 15-OCT-1999; 99US-0159783.

XX PA (CONJ-) CONJUCHEM INC.

XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX DR WPI; 2001-112059/12.

XX PT Modifying and attaching therapeutic peptides to albumin prevents  
XX PT peptidase degradation, useful for increasing length of in vivo activity

XX PS Disclosure; Page 589; 733pp; English.

XX CC The present invention describes a modified therapeutic peptide (I)  
XX CC comprising a therapeutically active amino acid region (III) and a  
XX CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
XX CC a less therapeutically active amino acid region (IV), which covalently  
XX CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX CC factors and neurotransmitters, to protect them from peptidase activity  
XX CC in vivo for the treatment of various disorders. Endogenous therapeutic  
XX CC peptides are not suitable as drug candidates as they require frequent  
XX CC administration due to rapid degradation by peptidases in the body.  
XX CC Modifying and attaching therapeutic peptides to albumin prevents or  
XX CC reduces the action of peptidases to increase length of activity (half  
XX CC life) and specificity as bonding to large molecules decreases  
XX CC intracellular uptake and interference with physiological processes.  
XX CC AAB90829 to AAB92441 represent peptides which can be used in the  
XX CC exemplification of the present invention.

XX SQ Sequence 24 AA;

Query Match 95.8%; Score 68; DB 22; Length 24;  
Best Local Similarity 90.9%; Pred. No. 0.00051;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPKPQWFWLM 11

Db 14 rpkpqwfwl1 24  
|||||||

RESULT 20

AAW09004

ID AAW09004 standard; peptide; 11 AA.

XX AC AAW09004;

XX DT 03-MAR-1997 (first entry)

XX DE Spantide analogue, acts as substance P antagonist.

XX KW Analogue; substance P; spantide; non-peptide bond;  
XX KW competitive inhibitor; receptor; neurogenic inflammation;  
XX KW rheumatoid arthritis; ulcerative colitis; eczema; Crohn's disease;  
XX KW anti-proliferative agent; small cell lung carcinoma; fibroblast.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT FT Misc-difference 1 /note= "D-form residue"

FT FT Modified-site 6..7

FT FT /label= Gln-psi[CH2-NH]-Trp

FT FT /note= "Opt. non-peptide bond, Claim 7"

FT FT Misc-difference 7

FT FT /note= "D-form residue"

FT FT Modified-site 7..8

FT FT /label= Trp-psi[CH2-NH]-Phe

FT FT /note= "Opt. non-peptide bond, Claim 6"

FT FT Modified-site 8..9

FT FT /label= Phe-psi[CH2-NH]-Trp

FT FT /note= "Opt. non-peptide bond"

FT FT Misc-difference 9

FT FT /note= "D-form residue"

FT FT Modified-site 9..10

FT FT /label= Trp-psi[CH2-NH]-Leu

FT FT /note= "Opt. non-peptide bond, Claim 4"

FT FT Modified-site 10..11

FT FT /label= Leu-psi[CH2-NH]-Nle

FT FT /note= "Opt. non-peptide bond, Claim 5"

FT FT Modified-site 11

FT FT /label= Nle

FT FT /note= "Amidated C-terminal"

XX US5410019-A.

XX PD 25-APR-1995.

XX PF 24-SEP-1987; 87US-0100571.

XX PR 30-MAR-1992; 92US-0860675.

XX PR 24-SEP-1987; 87US-0100571.

XX PR 25-MAR-1988; 88US-0173311.

XX PR 08-JUN-1988; 88US-0204171.

XX PR 16-JUN-1988; 88US-0207759.

XX PR 23-SEP-1988; 88US-0248771.

XX PR 14-OCT-1988; 88US-0257998.

XX PR 09-DEC-1988; 88US-0282328.

XX PR 02-MAR-1989; 89US-0317941.

XX PR 16-AUG-1989; 89US-0394727.

XX (TULA ) TULANE EDUCATIONAL FUND.

XX PI Coy DH, Moreau J;

XX DR WPI; 1995-169633/22.

XX PT Novel linear peptide substance P analogues - useful as substance P  
XX PT antagonists, for treating neurogenic inflammation

XX PS Claim 4-7; Column 20; 16pp; English.

XX CC The sequences given in AAW09003-04 represent analogues of substance P  
XX CC and spantide, respectively. These analogues comprise a non-peptide  
XX CC bond between an amino acid residue of the active site, which occurs

CC in the C-terminal half of the peptide, and an adjacent amino acid  
CC residue. They act as competitive inhibitors of the naturally  
CC occurring peptide by binding to its receptor. These peptides may be  
CC used in the treatment of diseases involving neurogenic inflammation,  
CC e.g. rheumatoid arthritis, ulcerative colitis, eczema and Crohn's  
CC disease. They are also useful as anti-proliferative agents, in  
CC the treatment of small cell lung carcinoma or disorders involving the  
CC proliferation of fibroblasts.

XX Sequence 11 AA;

Query Match 93.0%; Score 66; DB 16; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00048;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFwl 10

Db 1 rpkpqgfwl 10

RESULT 21

AAW99690

ID AAW99690 standard; peptide; 11 AA.

XX AC

XX AAW99690;

DT 03-JUN-1999 (first entry)

XX DE

XX Substance P analogue #7.

XX Substances P receptor antagonist; analgesic; inhibitor; NMDA blocker;  
KW nontoxic N-methyl-D-aspartate receptor antagonist; muscular pain;  
KW musculoskeletal pain; chronic pain; neuropathic pain; migraine.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Misc-difference 7 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 9..10 /note= "D-form residue"

FT Modified-site 11 /note= "Trp-psi(CH2-NH)-Leu"

FT /label= Nle

FT /note= "Norleucine, amidated"

XX WO9907413-A1.

XX 18-FEB-1999.

XX 26-MAY-1998; 98WO-US10707.

XX 11-AUG-1997; 97US-0055233.

XX (ALGO-) ALGOS PHARM CORP.

XX Caruso FS;

XX WPI; 1999-167216/14.

XX New analgesic composition comprises - a substance P receptor

PT antagonist with a substance P receptor antagonist potentiator, used

PT for the treatment of pain

XX Claim 3; Page 29; 54pp; English.

XX A method has been developed for treating pain with: (a) a substance P

CC receptor antagonist; and (b) a substance P receptor antagonist

CC potentiator, i.e. N-methyl-D-aspartate (NMDA) receptor antagonist or

CC substance that blocks at least 1 major intracellular consequence of  
CC NMDA receptor activation. The method can be used for treating muscular,  
CC musculoskeletal, chronic or neuropathic pain, or migraine. The present  
CC sequence represents a substance P analogue for use in the method.

XX Sequence 11 AA;

Query Match 93.0%; Score 66; DB 20; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00048;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFwl 10

Db 1 rpkpqgfwl 10

RESULT 22

AAW99691

ID AAW99691 standard; peptide; 11 AA.

XX AC

XX AAW99691;

XX 03-JUN-1999 (first entry)

XX DE

XX Substance P analogue #8.

XX Substances P receptor antagonist; analgesic; inhibitor; NMDA blocker;  
KW nontoxic N-methyl-D-aspartate receptor antagonist; muscular pain;  
KW musculoskeletal pain; chronic pain; neuropathic pain; migraine.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Misc-difference 7 /note= "D-form residue"

FT Modified-site 7..8 /note= "Trp-psi(CH2-NH)-Phe"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /label= Nle

FT /note= "Norleucine, amidated"

XX WO9907413-A1.

XX 18-FEB-1999.

XX 26-MAY-1998; 98WO-US10707.

XX 11-AUG-1997; 97US-0055233.

XX (ALGO-) ALGOS PHARM CORP.

XX Caruso FS;

XX WPI; 1999-167216/14.

XX New analgesic composition comprises - a substance P receptor

PT antagonist with a substance P receptor antagonist potentiator, used

PT for the treatment of pain

XX Claim 3; Page 29; 54pp; English.

XX A method has been developed for treating pain with: (a) a substance P

CC receptor antagonist; and (b) a substance P receptor antagonist

CC potentiator, i.e. N-methyl-D-aspartate (NMDA) receptor antagonist or

CC substance that blocks at least 1 major intracellular consequence of

CC NMDA receptor activation. The method can be used for treating muscular,

CC musculoskeletal, chronic or neuropathic pain, or migraine. The present

CC sequence represents a substance P analogue for use in the method.

XX  
SQ Sequence 11 AA;

Query Match 93.0%; Score 66; DB 20; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00048;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPOQWFWL 10  
DB 1 rpkpqgfwl 10  
|||||

RESULT 23  
AAP80315  
ID AAP80315 standard; protein; 11 AA.  
XX  
AC AAP80315;  
XX  
DT 14-SEP-1990 (first entry)  
XX  
DE Sequence of neuropeptide antagonist C which binds with polypeptide  
DE receptor for bombesin type polypeptides.  
XX  
KW Spantide; neuropeptide; polypeptide receptor; cancer diagnosis;  
KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides;  
KW antagonist C.  
XX  
OS Swiss 3T3 cells.  
XX

Key Location/Qualifiers  
FH Misc-difference 2 /label=OTHER  
FT /note="DPro"  
FT Misc-difference 7 /label=OTHER  
FT /note="DPhe"  
FT Misc-difference 1 /label=OTHER  
FT /note="DTrp"  
FT Misc-difference 11 /label=OTHER  
FT /note="Met-NH2"  
XX  
PN W08807551-A.  
XX  
PD 06-OCT-1988.  
XX  
PF 31-MAR-1988; 88WO-GB00255.  
XX  
PR 25-NOV-1987; 87GB-0027638.  
XX  
PA (IMCR ) IMPERIAL CANCER RES.  
XX  
PI Rosengurt E, Zachary I, Woll P;  
XX WPI; 1988-292842/41.  
XX  
PT New polypeptide receptor for bombesin type polypeptide(s) -  
PT is isolated from surface of Swiss 3T3 cells, and antibodies and  
PT antagonists are useful for treating uncontrolled cell proliferation  
XX  
PS Disclosure; Table 2; 42pp; English.  
XX

The patent claims a polypeptide isolated from the surface of Swiss 3T3  
CC cells which binds selectively with polypeptides of the bombesin type and  
CC binds with antagonist A and antagonist D. Antagonist A is a  
CC commercially available structural variant of substance P, known as  
CC [D-Arg1, D-Pro2, D-Trp7,9, Leu11] substance P. It is also known as  
CC [D-Pro2] spantide. Antagonist B is also commercially available structural  
CC variant of substance P, known as [D-Phe5] spantide. Substance P is an  
CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten  
CC substance P antagonists (see AAP80313-80322) were tested for their

ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue  
of bombesin in Swiss 3T3 cells). Antagonist D was clearly the most  
potent GRP antagonist. Peptides B, C, D, E, F, G, H, J and K were less  
potent than either A or D. Spantide (B) had no antagonist activity even  
at 100 uM. Polypeptide antagonists A and D and novel variants are useful  
for diagnosis and therapy, esp. of cancers where uncontrolled cell  
growth is associated with disorders of proteins of the bombesin family.

XX  
SQ Sequence 11 AA;

Query Match 85.9%; Score 61; DB 9; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.0028;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPOQWFWL 11  
DB 1 rpkpqgfwl 11  
|||||

RESULT 24  
AAW50968  
ID AAW50968 standard; peptide; 11 AA.  
XX  
AC AAW50968;  
XX  
DT 31-JUL-1998 (first entry)  
XX  
DE Substance P analogue, [D-Pro2,D-Phe7,D-Trp9].  
XX  
KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
KW Substance P; cancer; inhibition; growth hormone releasing factor;  
KW spantide.  
XX  
OS Synthetic.  
XX

Key Location/Qualifiers  
FH Misc-difference 2 /note="D-form residue"  
FT Misc-difference 7 /note="D-form residue"  
FT Misc-difference 9 /note="D-form residue"  
FT Modified-site 11 /note="C-terminal amide"  
XX  
PN EP835662-A2.  
XX  
PD 15-APR-1998.  
XX  
PF 11-DEC-1996; 96EP-0309012.  
XX  
PR 08-OCT-1996; 96US-0727679.  
PR 16-AUG-1996; 96IN-0001822.  
XX  
PA (NATM-) NAT INST IMMUNOLOGY.  
XX  
PI Jaggi M, Mukherjee R;  
XX WPI; 1998-208959/19.  
XX  
PT Composition containing analogues of vasoactive intestinal peptide,  
PT somatostatin - bombesin and substance P, for treatment of tumours  
PT and for inhibiting over-expression of these peptide(s)  
XX  
PS Disclosure; Page 13; 49pp; English.  
XX

The invention relates to a new composition which comprises: (i) the  
CC somatostatin analogue SOM2 AGCKNFQWKPTSDC (3-14 disulphide bridge),  
CC and (ii) at least 4 of the peptides: antagonist of vasoactive  
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP  
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin  
CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are

CC more general compositions containing peptide analogues of somatostatin,  
CC VIP, bombesin and substance P. The compositions are used in human or  
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour  
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,  
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,  
CC breast, kidney or particularly rectum and colon, and (b) to prevent,  
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
CC cells express receptors for VIP, somatostatin, bombesin and/or substance  
CC P. The present sequence represents a substance P analogue.  
XX  
SQ Sequence 11 AA;

Query Match 85.9%; Score 61; DB 19; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.0028;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
| | | | | : | | | |  
Db 1 rpkpqgffwlm 11

## RESULT 25

AAB98879  
ID AAB98879 standard; Peptide; 11 AA.

XX AAB98879;

DT 14-AUG-2001 (first entry)

DE Chimeric analgesic peptide #35.

XX Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 2 /note= "D-form residue"

FT Misc-difference 7 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11

FT /label= OTHER

FT /note= "C-terminal amide"

XX WO200130371-A2.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-US29789.

XX 28-OCT-1999; 99US-0428692.

XX (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;

XX WPI; 2001-397593/42.

XX New chimeric peptides used for treating pain comprise opioid receptor

XX binding group and nociceptive receptor binding group  
XX Claim 10; Page 15; 34pp; English.  
XX The present invention describes a number of chimeric peptides comprising  
XX an opioid receptor binding moiety and a nociceptive receptor binding  
XX moiety. These can be used as analgesics for the treatment of pain. Unlike  
XX opioid receptor based peptides alone, tolerance does not result from  
XX their long-term use. The present sequence is one of the peptides of the  
XX invention.

XX SQ Sequence 11 AA;

Query Match 85.9%; Score 61; DB 22; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.0028;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
| | | | | : | | | |  
Db 1 rpkpqgffwlm 11

## RESULT 26

AAB91412

ID AAB91412 standard; Peptide; 11 AA.

XX AAB91412;

DT 22-JUN-2001 (first entry)

DE Tachykinins peptide SEQ ID NO:588.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.

XX Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US13576.

XX 17-MAY-1999; 99US-0134406.

XX 10-SEP-1999; 99US-0153406.

XX 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents  
XX peptidase degradation, useful for increasing length of in vivo activity

XX Disclosure; Page 392; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)  
XX comprising a therapeutically active amino acid region (iii) and a  
XX reactive group (ii) (e.g. succinimidyl and maleimido groups) attached to  
XX a less therapeutically active amino acid region (iv), which covalently  
XX bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX factors and neurotransmitters, to protect them from peptidase activity  
XX in vivo for the treatment of various disorders. Endogenous therapeutic  
XX peptides are not suitable as drug candidates as they require frequent  
XX administration due to rapid degradation by peptidases in the body.  
XX Modifying and attaching therapeutic peptides to albumin prevents or  
XX reduces the action of peptidases to increase length of activity (half  
XX life) and specificity as bonding to large molecules decreases  
XX intracellular uptake and interference with physiological processes.  
XX AAB90829 to AAB92441 represent peptides which can be used in the  
XX exemplification of the present invention.

XX SQ Sequence 11 AA;

Query Match 85.9%; Score 61; DB 22; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.0028;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:||||  
Db 1 rpkpqgffwlm 11

## RESULT 27

AAB98880  
ID AAB98880 standard; Peptide; 12 AA.  
XX AC AAB98880;  
XX DT 14-AUG-2001 (first entry)  
XX DE  
XX DE Chimeric analgesic peptide #36.  
XX KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Misc-difference 2 /note= "D-form residue"  
FT Misc-difference 7 /note= "D-form residue"  
FT Misc-difference 9 /note= "D-form residue"  
FT Modified-site 12 /label= OTHER  
FT /note= "C-terminal amide"

WO200130371-A2.  
XX PN  
XX PD 03-MAY-2001.  
XX PF 27-OCT-2000; 2000WO-US29789.  
XX PR 28-OCT-1999; 99US-0428692.  
XX PA (NEW)- NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX DR WPI; 2001-397593/42.  
XX PT New chimeric peptides used for treating pain comprise opioid receptor  
XX binding group and nociceptive receptor binding group -  
XX PS Claim 10; Page 15-16; 34pp; English.

XX CC The present invention describes a number of chimeric peptides comprising  
CC an opioid receptor binding moiety and a nociceptive receptor binding  
CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
CC opioid receptor based peptides alone, tolerance does not result from  
CC their long-term use. The present sequence is one of the peptides of the  
CC invention.

XX SQ Sequence 12 AA;

Query Match 85.9%; Score 61; DB 22; Length 12;  
Best Local Similarity 90.9%; Pred. No. 0.003;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:||||  
Db 1 rpkpqgffwlm 11

## RESULT 28

AAP40481  
ID AAP40481 standard; Protein; 11 AA.  
XX AC AAP40481;  
XX DT 27-NOV-1991 (first entry)  
XX DE Substance P analogue.  
XX DE Substance P; analogue; antiinflammatory agent; analgesic.  
XX KW  
XX PN US4481139-A.  
XX PD 06-NOV-1984.  
XX PF 13-APR-1983; 83US-0484646.  
XX PR 13-APR-1983; 83US-0484646.  
XX PA (UYTE-) UNIVERSITY OF TEXAS SYSTEM.

XX PI Folkers K, Ji-cheng X;  
XX DR WPI; 1984-294258/47.  
XX FT Peptide analogues of substance P - useful as antagonists, e.g. as  
FT antiinflammatory agents and analgesics.

XX PS Claim 4; page 5; 5pp; English.

XX CC The peptide is a D-Arg1, D-Pro2, D-Phe 5, D-Trp7, D-Trp9, Leu11  
CC analogue of substance P. The peptide is a substance P antagonist  
CC with higher activity than known substance P analogues. It may be  
CC used as a biological research tool, ophthalmological antiinflammatory  
CC agent and analgesic.  
XX SQ Sequence 11 AA;

Query Match 84.5%; Score 60; DB 5; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.0039;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:||||  
Db 1 rpkpqgffwlm 11

## RESULT 29

AAP80316  
ID AAP80316 standard; protein; 11 AA.  
XX AC AAP80316;  
XX DT 14-SEP-1990 (first entry)

XX DE Sequence of neuropeptide antagonist D which binds with polypeptide  
DE receptor for bombesin type polypeptides.  
XX KW Spantide; neuropeptide; polypeptide receptor; cancer diagnosis;  
KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides.

XX OS Swiss 3T3 cells.

XX FH Key Location/Qualifiers  
FT Misc-difference 1 /label=OTHER  
FT /note="DArg"

FT Misc-difference 5 /label=OTHER  
FT /note="DPhe"

FT Misc-difference 7

FT /label=OTHER  
 FT Misc-difference 9 /note="D-Trp"  
 FT /label=OTHER  
 FT /note="D-Trp"  
 FT Misc-difference 11  
 FT /label=OTHER  
 FT /note="Leu-NH2"  
 PN W08807551-A.  
 XX  
 XX  
 PD 06-OCT-1988.  
 XX  
 XX 31-MAR-1988; 88WO-GB00255.  
 XX  
 XX 25-NOV-1987; 87GB-0027638.  
 XX  
 XX (IMCR ) IMPERIAL CANCER RES.  
 XX  
 XX Rosengurt E, Zachary I, Woll P;  
 XX  
 XX WPI; 1988-292842/41.  
 DR  
 XX

XX New polypeptide receptor for bombesin type polypeptide(s) -  
 PT is isolated from surface of Swiss 3T3 cells, and antibodies and  
 PT antagonists are useful for treating uncontrolled cell proliferation  
 XX  
 XX Disclosure; Table 2; 42pp; English.

XX The patent claims a polypeptide isolated from the surface of Swiss 3T3  
 CC cells which binds selectively with polypeptides of the bombesin type and  
 CC binds with antagonist A and antagonist D. Antagonist A is a  
 CC commercially available structural variant of substance P, known as  
 CC [D-Arg1, D-Pro2, D-Trp7,9, Leu11] substance P. It is also known as  
 CC [D-Pro2] spantide. Antagonist B is also commercially available structural  
 CC variant of substance P, known as [D-Phe5] spantide. Substance P is an  
 CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten  
 CC substance P antagonists (see AAP80313-80322) were tested for their  
 CC ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue  
 CC of bombesin in Swiss 3T3 cells). Antagonist D was clearly the most  
 CC potent GRP antagonist. Peptides B, C, D, E, F, G, H, J and K were less  
 CC potent than either A or D. Spantide (B) had no antagonist activity even  
 CC at 100 uM. Polypeptide antagonists A and D and novel variants are useful  
 CC for diagnosis and therapy, esp. of cancers where uncontrolled cell  
 CC growth is associated with disorders of proteins of the bombesin family.

XX Sequence 11 AA;

Query Match 84.5%; Score 60; DB 9; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.0039;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQQWFNLM 11  
 |||||  
 Db 1 rpkpgfwll 11

RESULT 30

AAW50979  
 ID AAW50979 standard; peptide; 11 AA.  
 XX  
 AC AAW50979;

31-JUL-1998 (first entry)

Substance P analogue [D-Trp2,7,9].

Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
 Substante P; cancer; inhibition; growth hormone releasing factor;  
 spantide.  
 Synthetic.

XX Key Location/Qualifiers  
 FT Misc-difference 2 /note= "D-form residue"  
 FT  
 FT Misc-difference 7 /note= "D-form residue"  
 FT  
 FT Misc-difference 9 /note= "D-form residue"  
 FT  
 FT Modified-site 11 /note= "D-form residue"  
 FT  
 FT  
 XX  
 XX EP835662-A2.  
 PN  
 XX  
 PD 15-APR-1998.  
 XX  
 XX 11-DEC-1996; 96EP-0309012.  
 PF  
 XX 08-OCT-1996; 96US-0727679.  
 PR  
 PR 16-AUG-1996; 96IN-0001822.  
 XX  
 XX (NAIM-) NAT INST IMMUNOLOGY.  
 PA  
 XX Jaggi M, Mukherjee R;  
 PI  
 XX WPI; 1998-208959/19.  
 DR  
 XX  
 XX Composition containing analogues of vasoactive intestinal peptide,  
 PT somatostatin - bombesin and substance P, for treatment of tumours  
 PT and for inhibiting over-expression of these peptide(s)  
 PT  
 XX Disclosure; Page 13; 49pp; English.

XX The invention relates to a new composition which comprises: (i) the  
 CC somatostatin analogue SOM2 AGCKNFTdWKPTSDC (3-14 disulphide bridge),  
 CC and (ii) at least 4 of the peptides: antagonist of vasoactive  
 CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP  
 CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin  
 CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are  
 CC more general compositions containing peptide analogues of somatostatin,  
 CC VIP, bombesin and substance P. The compositions are used in human or  
 CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour  
 CC or cancer cells, particularly for treatment of leukaemia, lymphoma,  
 CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,  
 CC breast, kidney or particularly rectum and colon, and (b) to prevent,  
 CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
 CC cells express receptors for VIP, somatostatin, bombesin and/or substance  
 CC P. The present sequence represents a substance P analogue.

XX Sequence 11 AA;

Query Match 84.5%; Score 60; DB 19; Length 11;  
 Best Local Similarity 90.9%; Pred. No. 0.0039;  
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQQWFNLM 11  
 | | | | |  
 Db 1 rwkpgqfwlm 11

RESULT 31

AAW50972  
 ID AAW50972 standard; peptide; 11 AA.  
 XX  
 AC AAW50972;

31-JUL-1998 (first entry)

Substance P analogue, [D-Arg1,D-Phe5,D-Trp7,9,Leu11].

Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
 Substante P; cancer; inhibition; growth hormone releasing factor;  
 spantide.



```

XX OS Synthetic.
XX KW Key Location/Qualifiers
XX FH Misc-difference 1 /note= "D-form residue"
XX FT Misc-difference 5 /note= "D-form residue"
XX FT Misc-difference 7 /note= "D-form residue"
XX FT Misc-difference 9 /note= "D-form residue"
XX FT Misc-difference 11 /note= "D-form residue"
XX FT Modified-site 13 /note= "C-terminal amide"
XX PN EP835662-A2.
XX PD 15-APR-1998.
XX PF 11-DEC-1996; 96EP-0309012.
XX PR 08-OCT-1996; 96US-0727679.
XX PR 16-AUG-1996; 96IN-0001822.
XX PA (NAIM-) NAT INST IMMUNOLOGY.
XX PI Jaggi M, Mukherjee R;
XX DR WPI; 1998-208959/19.
XX PT Composition containing analogues of vasoactive intestinal peptide,
XX PT somatostatin - bombesin and substance P, for treatment of tumours
XX PT and for inhibiting over-expression of these peptide(s)
XX PS Disclosure; Page 13; 49pp; English.
XX CC The invention relates to a new composition which comprises: (i) the
XX CC somatostatin analogue SOM2 AGCKNFDWKPTSDC (3-14 disulphide bridge),
XX CC and (ii) at least 4 of the peptides: antagonist of vasoactive
XX CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP
XX CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin
XX CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are
XX CC more general compositions containing peptide analogues of somatostatin,
XX CC VIP, bombesin and substance P. The compositions are used in human or
XX CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour
XX CC or cancer cells, particularly for treatment of leukaemia, lymphoma,
XX CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,
XX CC breast, kidney or particularly rectum and colon, and (b) to prevent,
XX CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer
XX CC cells express receptors for VIP, somatostatin, bombesin and/or substance
XX CC P. The present sequence represents a substance P analogue.
XX SQ Sequence 11 AA;

Query Match 84.5%; Score 60; DB 19; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.0039;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11
Db 1 rpkipqfwll 11

RESULT 32
AAW50942
ID AAW50942 standard; peptide; 11 AA.
XX AC AAW50942;
XX DT 31-JUL-1998 (first entry)
XX DE Substance P antagonist (SP1).

```

```

XX Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;
XX KW Substance P; cancer; inhibition.
XX OS Synthetic.
XX FH Key Location/Qualifiers
XX FT Misc-difference 1 /note= "D-form residue"
XX FT Misc-difference 5 /note= "D-form residue"
XX FT Misc-difference 7 /note= "D-form residue"
XX FT Misc-difference 9 /note= "D-form residue"
XX FT Modified-site 13 /note= "C-terminal amide"
XX PN EP835662-A2.
XX PD 15-APR-1998.
XX PF 11-DEC-1996; 96EP-0309012.
XX PR 08-OCT-1996; 96US-0727679.
XX PR 16-AUG-1996; 96IN-0001822.
XX PA (NAIM-) NAT INST IMMUNOLOGY.
XX PI Jaggi M, Mukherjee R;
XX DR WPI; 1998-208959/19.
XX PT Composition containing analogues of vasoactive intestinal peptide,
XX PT somatostatin - bombesin and substance P, for treatment of tumours
XX PT and for inhibiting over-expression of these peptide(s)
XX PS Claim 1; Page 4; 49pp; English.
XX CC The invention relates to a new composition which comprises: (i) the
XX CC somatostatin analogue SOM2 AGCKNFDWKPTSDC (3-14 disulphide bridge),
XX CC and (ii) at least 4 of the peptides: antagonist of vasoactive
XX CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP
XX CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin
XX CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are
XX CC more general compositions containing peptide analogues of somatostatin,
XX CC VIP, bombesin and substance P. The compositions are used in human or
XX CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour
XX CC or cancer cells, particularly for treatment of leukaemia, lymphoma,
XX CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,
XX CC breast, kidney or particularly rectum and colon, and (b) to prevent,
XX CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer
XX CC cells express receptors for VIP, somatostatin, bombesin and/or substance
XX CC P. The present sequence represents substance P antagonist (SP1).
XX SQ Sequence 11 AA;

Query Match 84.5%; Score 60; DB 19; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.0039;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11
Db 1 rpkipqfwll 11

RESULT 33
AAB08303
ID AAB08303 standard; peptide; 11 AA.
XX AC AAB08303;
XX DE

```

DT 04-DEC-2000 (first entry)  
XX Amino acid sequence of Substance P analogue SP1.  
DE  
XX  
XX Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;  
KW VIP1; VIP2; VIP3; BOM1; bombesin; SP1; substance P; MuJ-7; tumour growth;  
KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;  
KW leukaemia; lymphoma.  
XX  
XX Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1 /note= "D-form residue"  
FT Misc-difference 5 /note= "D-form residue"  
FT Misc-difference 7 /note= "D-form residue"  
FT Misc-difference 9 /note= "D-form residue"  
FT Misc-difference 9 /note= "D-form residue"  
XX  
XX WO200047221-A1.  
XX  
XX 17-AUG-2000.  
XX  
XX 11-FEB-2000; 2000WO-US03559.  
XX  
XX 11-FEB-1999; 99US-0248381.  
XX  
XX (NAIM-) NAT INST IMMUNOLOGY.  
PA (DABU-) DABUR RES FOUND.  
PA (CORD/) CORD J I.  
XX  
XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;  
PI Singh A;  
XX  
XX WPI; 2000-549083/50.  
XX  
XX Novel therapeutically active composition comprising at least 5  
PT peptides, useful for treating angiogenesis especially as a result of  
PT adenocarcinomas -  
XX  
XX Disclosure; Page 8; 42pp; English.  
XX  
XX The present sequence represents an analogue of Substance P. The  
CC specification describes therapeutically active compositions comprising  
CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and  
CC at least four analogues chosen from vasoactive intestinal peptide (VIP) 1  
CC (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a VIP  
CC receptor antagonist), BOM1 (a bombesin antagonist), and SP1 (a substance  
CC P antagonist). The combination of these 7 analogues is known as MuJ-7.  
CC MuJ-7 is used as an anticancer drug to restrict tumour growth and spread  
CC by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits  
CC metastasis through its antiangiogenic activity in all cancers. The  
CC peptides are useful for the treatment and prevention of angiogenesis,  
CC especially as a result of adenocarcinomas of the colon, breast, lung,  
CC prostate, kidney, leukemias or lymphomas.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 84.5%; Score 60; DB 21; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.0039;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RPKPQQWFWM 11  
Db 1 rpkpfgfwll 11  
RESULT 34  
AAB91414  
ID AAB91414 standard; Peptide; 11 AA.

XX AAB91414;  
AC  
XX 22-JUN-2001 (first entry)  
DT  
XX Tachykinins peptide SEQ ID NO:590.  
DE  
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
XX Homo sapiens.  
OS Synthetic.  
OS  
XX WO200069900-A2.  
PN  
XX 23-NOV-2000.  
PD  
XX 17-MAY-2000; 2000WO-US13576.  
PF  
XX 17-MAY-1999; 99US-0134406.  
PR  
XX 10-SEP-1999; 99US-0153406.  
PR  
XX 15-OCT-1999; 99US-0159783.  
XX  
XX (CONJ-) CONJUCHEM INC.  
PA  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX WPI; 2001-112059/12.  
XX  
XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
XX Disclosure; Page 392; 733pp; English.  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 84.5%; Score 60; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.0039;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RPKPQQWFWM 11  
Db 1 rpkpfgfwll 11  
RESULT 35  
AAB08313  
ID AAB08313 standard; peptide; 11 AA.  
XX  
XX AAB08313;  
AC  
XX 04-DEC-2000 (first entry)  
DT

XX Amino acid sequence of an antiangiogenic peptide.

XX DE

XX KW Vasoactive intestinal peptide; VIP; analogue; somatostatin; SOM1; SOM2;

XX KW VIP1; VIP2; VIP3; BOM1; bombesin; SPI; substance P; MuJ-7; tumour growth;

XX KW tumour angiogenesis; metastasis; cancer; angiogenesis; adenocarcinoma;

XX KW leukaemia; lymphoma.

XX OS

XX Synthetic.

XX FH

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Modified-site 5 /label= Aib

FT /note= "alpha-aminoisobutyric acid"

FT Misc-difference 7 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 10 /label= Aib

FT /note= "alpha-aminoisobutyric acid"

FT Modified-site 11 /note= "amidated residue"

XX KW

XX WO200047221-A1.

XX PD

XX 17-AUG-2000.

XX PF

XX 11-FEB-2000; 2000WO-US03559.

XX PR

XX 11-FEB-1999; 99US-0248381.

XX PA

XX (NAIM-) NAT INST IMMUNOLOGY.

XX PA (DABU-) DABUR RES FOUND.

XX PA (CORD/) CORD J I.

XX PI

XX Mukherjee R, Jaggi M, Prasad S, Burman AC, Rajendran P, Mathur A;

XX PI Singh AT;

XX DR

XX WPI; 2000-549083/50.

XX PS

XX Novel therapeutically active composition comprising at least 5

XX PT peptides, useful for treating angiogenesis especially as a result of

XX PT adenocarcinomas -

XX PS

XX Claim 11; Page 31; 42pp; English.

XX CC

XX AAB08304-15 represent peptides which have an antiangiogenic effect. The

XX CC specification describes therapeutically active compositions comprising

XX CC at least one analogue of somatostatin (chosen from SOM1 and SOM2), and

XX CC at least four analogues chosen from vasoactive intestinal peptide (VIP)

XX CC 1 (a VIP antagonist), VIP2 (a VIP receptor binding inhibitor), VIP3 (a

XX CC VIP receptor antagonist), BOM1 (a bombesin antagonist), and SPI (a

XX CC substance P antagonist). The combination of these 7 analogues is known as

XX CC MuJ-7. MuJ-7 is used as an anticancer drug to restrict tumour growth and

XX CC spread by inhibiting tumour angiogenesis. MuJ-7, in addition, inhibits

XX CC metastasis through its antiangiogenic activity in all cancers. The

XX CC peptides are useful for the treatment and prevention of angiogenesis,

XX CC especially as a result of adenocarcinomas of the colon, breast, lung,

XX CC prostate, kidney, leukemias or lymphomas.

XX SQ

XX Sequence 11 AA;

Query Match 80.3%; Score 57; DB 21; Length 11;

Best Local Similarity 72.7%; Pred. No. 0.011;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11

Db 1 rpkpxqfwf 11

RESULT 36

AAP80322

ID AAP80322 standard; protein; 8 AA.

XX AC

XX AAP80322;

XX DT

XX 14-SEP-1990 (first entry)

XX DE

XX Sequence of neuropeptide antagonist K which binds with polypeptide

XX DE receptor for bombesin type polypeptides.

XX KW

XX Spantide; neuropeptide; polypeptide receptor; cancer diagnosis;

XX KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides;

XX KW antagonist K.

XX OS

XX Swiss 3T3 cells.

XX FH

XX Key Location/Qualifiers

FT Misc-difference 1 /label=OTHER

FT /note= "DPro"

FT Misc-difference 4 /label=OTHER

FT /note= "DTrip"

FT Misc-difference 6 /label=OTHER

FT /note= "DTrip"

FT Misc-difference 8 /label=OTHER

FT /note= "DTrip"

FT Misc-difference 8 /label=OTHER

FT /note= "Met-NH2"

XX PN

XX WO8807551-A.

XX XX

XX 06-OCT-1988.

XX XX

XX 31-MAR-1988; 88WO-GB00255.

XX PR

XX 25-NOV-1987; 87GB-0027638.

XX XX

XX (IMCR ) IMPERIAL CANCER RES.

XX XX

XX Rosengurt E, Zachary I, Woll P;

XX PI

XX WPI; 1988-292842/41.

XX DR

XX New polypeptide receptor for bombesin type polypeptide(s) -

XX PT is isolated from surface of Swiss 3T3 cells, and antibodies and

XX PT antagonists are useful for treating uncontrolled cell proliferation

XX PS

XX Disclosure; Table 2; 42pp; English.

XX CC

XX The patent claims a polypeptide isolated from the surface of Swiss 3T3

XX CC cells which binds selectively with polypeptides of the bombesin type and

XX CC binds with antagonist A and antagonist D. Antagonist A is a

XX CC commercially available structural variant of substance P, known as

XX CC [D-Arg1, D-Pro2, D-Trip7,9, Leu11] substance P. It is also known as

XX CC [D-Pro2] spantide. Antagonist B is also commercially available structural

XX CC variant of substance P, known as [D-Phe5] spantide. Substance P is an

XX CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten

XX CC substance P antagonists (see AAP80313-80322) were tested for their

XX CC ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue

XX CC of bombesin in Swiss 3T3 cells). Antagonist D was clearly the most

XX CC potent GRP antagonist. Peptides B, C, D, E, F, G, H, J and K were less

XX CC potent than either A or D. Spantide (B) had no antagonist activity even

XX CC at 100 uM. Polypeptide antagonists A and D and novel variants are useful

XX CC for diagnosis and therapy, esp. of cancers where uncontrolled cell

XX CC growth is associated with disorders of proteins of the bombesin family.

XX SQ

XX Sequence 8 AA;

Query Match 76.1%; Score 54; DB 9; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.3e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PQQWFWM 11  
| | | | | | | |  
Db 1 pqqwfwlm 8

## RESULT 37

AAW50970  
ID AAW50970 standard; peptide; 8 AA.

XX AC AAW50970;

XX DT 31-JUL-1998 (first entry)

XX DE Substance P analogue residues 4-11, [D-Pro4,D-Trp7,9].

XX KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
XX KW Substance P; cancer; inhibition; growth hormone releasing factor;  
XX KW spantide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Misc-difference 4 /note= "D-form residue"

FT Misc-difference 6 /note= "D-form residue"

FT Modified-site 8 /note= "C-terminal amide"

XX EP835662-A2.

XX 15-APR-1998.

XX 11-DEC-1996; 96EP-0309012.

XX 08-OCT-1996; 96US-0727679.

XX 16-AUG-1996; 96IN-0001822.

XX (NAIM-) NAT INST IMMUNOLOGY.

XX Jaggi M, Mukherjee R;

XX WPI; 1998-208959/19.

XX Composition containing analogues of vasoactive intestinal peptide,  
XX somatostatin - bombesin and substance P, for treatment of tumours  
XX and for inhibiting over-expression of these peptide(s)

XX Disclosure; Page 13; 49pp; English.

XX The invention relates to a new composition which comprises: (i) the  
XX somatostatin analogue SOM2 ACCKNFFQWKPTSC (3-14 disulphide bridge),  
XX and (ii) at least 4 of the peptides: antagonist of vasoactive  
XX intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP  
XX receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin  
XX antagonist (BOM1) and substance P antagonist (SP1). Also claimed are  
XX more general compositions containing peptide analogues of somatostatin,  
XX VIP, bombesin and substance P. The compositions are used in human or  
XX veterinary medicine: (a) to kill (or inhibit multiplication of) tumour  
XX or cancer cells, particularly for treatment of leukaemia, lymphoma,  
XX adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,  
XX breast, kidney or particularly rectum and colon, and (b) to prevent,  
XX inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
XX cells express receptors for VIP, somatostatin, bombesin and/or substance  
XX P. The present sequence represents a substance P analogue.

XX Sequence 8 AA;

Query Match 76.1%; Score 54; DB 19; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.3e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PQQWFWM 11  
| | | | | | | |  
Db 1 pqqwfwlm 8

## RESULT 38

AAW50970

ID AAR28392 standard; peptide; 11 AA.

XX AC AAR28392;

XX DT 18-MAR-1993 (first entry)

XX DE Bradykinin receptor antagonist CT-0008.

XX KW Bradykinin receptor antagonist; heterodimer; higher oligomer;  
XX KW potency; duration; CP-0088; burns; migraine; shock CNS injury; asthma;  
XX KW rhinitis; premature labour; inflammatory arthritis; homodimer;  
XX KW inflammatory bowel disease.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Misc-difference 2 /note= "D-form residue"

FT Misc-difference 7 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

FT Modified-site 11 /label= Nle

XX WO9217201-A.

XX 15-OCT-1992.

XX 30-MAR-1992; 92WO-US02431.

XX 01-APR-1991; 91US-0677391.

XX 27-MAR-1992; 92US-0859582.

XX (CORT-) CORTECH INC.

XX Allen LG, Blodgett JK, Cheronis JC, Eubanks SR, Nguyen KT;

XX Whalley ET;

XX WPI; 1992-365995/44.

XX Bradykinin antagonists comprising linked bradykinin antagonist  
XX chains - are for treatment of post-operative pain, asthma and  
XX aseptic shock

XX Disclosure; Page 76; 109pp; English.

XX The sequence given is a bradykinin receptor antagonist which can form  
XX homo- or heterodimers or higher oligomers. It demonstrates greater  
XX potency and/or duration of action than the parent peptide itself.  
XX Bradykinin receptors antagonists such as this can be used in the  
XX treatment of burns, peroperative pain, migraine and other forms of  
XX pain, shock CNS injury, asthma, rhinitis, premature labour,  
XX inflammatory arthritis, inflammatory bowel disease etc.

XX Sequence 11 AA;

Query Match 71.8%; Score 51; DB 13; Length 11;  
Best Local Similarity 80.0%; Pred. No. 0.089;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQOWFWL 10  
| | | | | : | | | |  
Db 1 rpkpqffwl 10

## RESULT 39

AAP30141  
ID AAP30141 standard; peptide; 11 AA.

XX AC AAP30141;

XX 14-JUN-1992 (first entry)

DE Sequence of peptide with substance P inhibiting activity.

XX Substance P antagonist; pain therapy; hypertension.

XX Key Location/Qualifiers

FT Modified-site 2 /label= D-P, D-p-Cl-F

FT Modified-site 7 /label= D-W

FT Misc-difference 8 /label= F, I

FT Modified-site 9 /label= D-W

FT Modified-site 11 /label= M, I

FT /note= "bonded to NH2"

XX W08301251-A.

PN 14-APR-1983.

XX 09-OCT-1981; 81WO-DE00171.

XX 09-OCT-1981; 81WO-DE00171.

PR 09-OCT-1981; 81EP-0902802.

XX (FERR ) FERRING ARZNEIMITTE.

PA (HORI/) HORIG J.

XX Horig J;

XX WPI; 1983-39155K/16 (39155K).

XX Undeca; peptide derivs. with substance P inhibiting activity -

PT useful for treating pain and hypertension

PS Claim 1; Page 18; 25pp; German.

XX The peptides of the invention are powerful antagonists of Substance

CC P and so are useful in human and veterinary medicine, for treating

CC pain and hypertension (esp.) chronic conditions. A 10 microm concn.

CC of AAP30142 produced about 50% inhibition at a Substance P concn. of

CC 7.5-20 nanom.

XX Sequence 11 AA;

Query Match 70.4%; Score 50; DB 4; Length 11;  
Best Local Similarity 80.0%; Pred. No. 0.13;  
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 RPKPQOWFWL 10  
| | | | | : | | | |  
Db 1 rxxpqgwxwl 10

OS Synthetic.  
XX Key Location/Qualifiers  
FH Modified-site 9  
FT /label= MeGly  
FT Modified-site 11  
FT /label= OTHER  
FT /note= "Met(O)2-NH2"  
XX WO9218536-A.  
XX  
XX 29-OCT-1992.  
XX  
XX 22-APR-1992; 92WO-US03307.  
XX  
XX 22-APR-1991; 91EP-0200955.  
XX  
XX (MIGW ) MALLINCKRODT MEDICAL INC.  
XX  
XX Bakker WH, Hagen PM, Krenning EP, Lamberts SWJ, Visser TJ;  
XX WPI; 1992-382047/46.  
XX  
XX Detection and localisation of tissues with neurokinine-1 receptors  
PT - for detecting and treating tumours having neurokinine-1  
PT receptors e.g. malignant glioma, small cell lung cancer etc.  
XX  
XX Disclosure; Page 4; 22pp; English.  
XX  
XX This peptide or its Tyr0 deriv. is a preferred peptide having a  
CC selective affinity to neurokinine-1 receptors which (when  
CC labelled with a radioactive isotope) can be used in imaging methods.  
CC A generic formula for preferred peptides is AAR28441. Such peptides  
CC are thus useful in diagnosis and treatment of conditions that are  
CC related to NK1 receptors and in visualising NK1 receptors on certain  
CC tissues. See AAR28442-R28446.  
XX  
XX Sequence 11 AA;  
SQ  
  
Query Match 69.0%; Score 49; DB 13; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.18;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPKPQQWFWM 11  
DB 1 rpkpqffmlm 11  
  
RESULT 42  
AAW92666  
ID AAW92666 standard; peptide; 11 AA.  
XX  
AC AAW92666;  
XX  
DT 30-APR-1999 (first entry)  
XX  
DE Human tachykinin agonist beta-amyloid peptide fragment #12.  
XX  
KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congenital angiodysplasia.  
XX  
XX Homo sapiens.  
OS  
XX US5876948-A.  
XX  
XX 02-MAR-1999.  
XX  
XX 27-JUL-1991; 91US-0737371.  
XX  
XX 29-JUL-1991; 91US-0737371.  
XX  
XX 27-JUL-1990; 90US-0559173.  
XX

XX (CHIL-) CHILDRENS MEDICAL CENT.  
XX  
XX Yankner BA;  
XX  
XX WPI; 1999-189630/16.  
XX  
XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
PT  
XX  
XX Disclosure; Column 15-16; 28pp; English.  
XX  
XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congenital angiodysplasia with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.  
XX  
XX Sequence 11 AA;  
SQ  
  
Query Match 69.0%; Score 49; DB 20; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.18;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPKPQQWFWM 11  
DB 1 rpkpqyfglm 11  
  
RESULT 43  
AAP80318  
ID AAP80318 standard; protein; 8 AA.  
XX  
AC AAP80318;  
XX  
DT 14-SEP-1990 (first entry)  
XX  
DE Sequence of neuropeptide antagonist F which binds with polypeptide  
DE receptor for bombesin type polypeptides.  
XX  
XX Spantide; neuropeptide; polypeptide receptor; cancer diagnosis;  
KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides;  
KW antagonist F.  
XX  
XX Swiss 3T3 cells.  
OS  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1 /label=OTHER  
FT /note="DPro"  
FT Misc-difference 4 /label=OTHER  
FT /note="DTrp"  
FT Misc-difference 6 /label=OTHER  
FT /note="DTrp"  
FT Misc-difference 7 /label=OTHER  
FT /note="DTrp"  
FT Misc-difference 8 /label=OTHER  
FT /note="DTrp"  
FT Misc-difference 8 /label=OTHER  
FT /note="Met-NH2"  
XX  
XX WO8807551-A.  
XX  
XX 06-OCT-1988.  
XX  
XX 31-MAR-1988; 88WO-GB00255.  
XX

XX 25-NOV-1987; 87GB-0027638.  
XX (IMCR ) IMPERIAL CANCER RES.  
XX Rosengurt E, Zachary I, Woll P;  
XX WPI; 1988-292842/41.  
XX  
XX New polypeptide receptor for bombesin type polypeptide(s) -  
PT is isolated from surface of Swiss 3T3 cells, and antibodies and  
PT antagonists are useful for treating uncontrolled cell proliferation  
XX  
XX Disclosure; Table 2; 42pp; English.  
XX  
XX The patent claims a polypeptide isolated from the surface of Swiss 3T3  
CC cells which binds selectively with polypeptides of the bombesin type and  
CC binds with antagonist A and antagonist D. Antagonist A is a  
CC commercially available structural variant of substance P, known as  
CC [D-Arg1, D-Pro2, D-Trp7,9, Leu11] substance P. It is also known as  
CC [D-Pro2] spantide. Antagonist B is also commercially available structural  
CC variant of substance P, known as [D-Phe5] spantide. Substance P is an  
CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten  
CC substance P antagonists (see AAP80313-80322) were tested for their  
CC ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue  
CC of bombesin in Swiss 3T3 cells). Antagonist D was clearly the most  
CC potent GRP antagonist. Peptides B, C, D, E, F, G, H, J and K were less  
CC potent than either A or D. Spantide (B) had no antagonist activity even  
CC at 100 uM. Polypeptide antagonists A and D and novel variants are useful  
CC for diagnosis and therapy, esp. of cancers where uncontrolled cell  
CC growth is associated with disorders of proteins of the bombesin family.  
XX  
XX Sequence 8 AA;  
SQ

Query Match 67.6%; Score 48; DB 9; Length 8;  
Best Local Similarity 87.5%; Pred. No. 4.3e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 PQQWFWM 11  
DB 1 pqqwfwmm 8

RESULT 44  
AAR15359  
ID AAR15359 standard; Protein; 8 AA.  
XX  
XX AAR15359;  
XX  
XX 02-MAR-1992 (first entry)  
XX  
XX Substance P antagonist (1).  
XX  
XX Analgesic; antiinflammatory; oxidation; resistance; sterilisation;  
KW central nervous system.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Modified-site 1 /label= D-Pro  
FT Modified-site 4 /label= D-Trp  
FT Modified-site 6 /label= D-Trp  
FT Modified-site 7 /label= D-Trp  
FT Modified-site 7 /label= D-ethyl-Trp  
XX  
XX WO9118016-A.  
XX  
XX 28-NOV-1991.  
XX

PF 17-APR-1991; 91WO-EP00727.  
XX  
XX 11-MAY-1990; 90IT-0020273.  
XX (DEGH/) DEGHENGI R.  
XX Deghenghi R;  
XX WPI; 1991-369186/50.  
XX  
XX New bioactive peptide(s) - having D-tryptophan replaced by  
PT D-2-alkyl-tryptophan to increase resistance to oxidative  
PT degradation  
XX  
XX Claim 9; Page 21; 26pp; English.  
XX  
XX The peptides represented in AAR15359 and AAR15360 are antagonists of  
CC substance P. Substance P is a neurotransmitter used by sensory  
CC neurons that convey responses of pain or other noxious stimuli to  
CC the central nervous system. These peptides have analgesic and  
CC antiinflammatory activity.  
CC The D-2-alkyl-Trp provides increased oxidation resistance to the  
CC peptide while maintaining the same pharmacological effect as  
CC analogous bioactive peptides in which the tryptophan residue  
CC is not replaced. Oxidative degradation may take place e.g in the  
CC presence of reactive radicals or during high energy sterilisation.  
CC See also AAR15357-63.  
XX  
XX Sequence 8 AA;  
SQ

Query Match 67.6%; Score 48; DB 12; Length 8;  
Best Local Similarity 87.5%; Pred. No. 4.3e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 PQQWFWM 11  
DB 1 pqqwfwmm 8

RESULT 45  
AAW50974  
ID AAW50974 standard; peptide; 8 AA.  
XX  
XX AAW50974;  
XX  
XX 31-JUL-1998 (first entry)  
XX  
XX Substance P analogue residues 4-11, [D-Pro4,D-Trp7,9,10].  
XX  
XX Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
KW Substance P; cancer; inhibition; growth hormone releasing factor;  
KW spantide.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 1 /note= "D-form residue"  
FT Misc-difference 4 /note= "D-form residue"  
FT Misc-difference 6 /note= "D-form residue"  
FT Misc-difference 7 /note= "D-form residue"  
FT Misc-difference 8 /note= "D-form residue"  
FT Modified-site 8 /note= "C-terminal amide"  
XX  
XX EP835662-A2.  
XX  
XX 15-APR-1998.  
XX  
XX 11-DEC-1996; 96EP-0309012.  
XX

XX 08-OCT-1996; 96US-0727679.  
PR 16-AUG-1996; 96IN-0001822.  
XX (NATM-) NAT INST IMMUNOLOGY.  
XX Jaggi M, Mukherjee R;  
PI WPI; 1998-208959/19.  
DR Composition containing analogues of vasoactive intestinal peptide,  
PT somatostatin - bombesin and substance P, for treatment of tumours  
PT and for inhibiting over-expression of these peptide(s)  
XX Disclosure; Page 13; 49pp; English.  
XX The invention relates to a new composition which comprises: (i) the  
CC somatostatin analogue SOM2 AGCKNFRDWRKPTSDC (3-14 disulphide bridge),  
CC and (ii) at least 4 of the peptides: antagonist of vasoactive  
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP  
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin  
CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are  
CC more general compositions containing peptide analogues of somatostatin,  
CC VIP, bombesin and substance P. The compositions are used in human or  
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour  
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,  
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,  
CC breast, kidney or particularly rectum and colon, and (b) to prevent,  
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
CC cells express receptors for VIP, somatostatin, bombesin and/or substance  
CC P. The present sequence represents a substance P analogue.  
XX Sequence 8 AA;  
SQ

Query Match 67.6%; Score 48; DB 19; Length 8;  
Best Local Similarity 87.5%; Pred. No. 4.3e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 4 PQQWFWM 11  
DB 1 pqqwfwm 8

RESULT 46  
AAP61480  
ID AAP61480 standard; peptide; 11 AA.  
AC AAP61480;  
XX 22-AUG-1991 (first entry)  
DT Sequence of undeca peptide substance P1.  
XX Hypertension therapy; sleep disorder; anti-stress agent.  
KW Key Location/Qualifiers  
FH Misc-difference 11  
FT /label= Met-NH2  
XX DD229593-A.  
XX 13-NOV-1985.  
XX 28-NOV-1984; 84DD-0269954.  
XX 28-NOV-1984; 84DD-0269954.  
XX (DEAK ) AKAD WISSENSCHAFT DDR.  
XX Oehme P, Hecht K, Wachtel E, Roske I, Kolometsewa IA;  
PI Alrapetjan M, Blenert M, Vogt WE, Hilse H, Gores E, Poppei M;  
PI Nieber K, Bergmann J;

XX WPI; 1986-069587/11.  
DR Cpd, having N-terminal sequences of undeca-peptide substance P -  
PT are medicinal agents with anti-stress activity  
XX Claim 1; Page 1; 15pp; German.  
XX The inventors claim an antistress compound which contains the N-  
CC terminal SQ of AAP61480, pref. Arg-Pro-Lys-Pro-X (X= COOH or NH2).  
CC Compared with the full undecapeptide they have much reduced  
CC side effects (acute hypotension, spastic effects on the ileum and  
CC histamine release from peritoneal mast cells).  
XX Sequence 11 AA;  
SQ

Query Match 67.6%; Score 48; DB 7; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQOWFWLM 11  
DB 1 rpkpqffglm 11

RESULT 47  
AAP80312  
ID AAP80312 standard; protein; 11 AA.  
AC AAP80312;  
XX 14-SEP-1990 (first entry)  
DT Sequence of neuropeptide substance P which binds with polypeptide  
DE receptor for bombesin type polypeptides.  
XX Spantide; neuropeptide; polypeptide receptor; cancer diagnosis;  
KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides;  
KW substance P.  
XX Swiss 3T3 cells.  
XX Key Location/Qualifiers  
FH Misc-difference 11  
FT /label=OTHER  
FT /note="Met-NH2"  
XX WO8807551-A.  
XX 06-OCT-1988.  
XX 31-MAR-1988; 88WO-GB00255.  
XX 25-NOV-1987; 87GB-0027638.  
XX (IMCR ) IMPERIAL CANCER RES.  
XX Rosengurt E, Zachary I, Woll P;  
PI WPI; 1988-292842/41.  
XX New polypeptide receptor for bombesin type polypeptide(s) -  
PT is isolated from surface of Swiss 3T3 cells, and antibodies and  
PT antagonists are useful for treating uncontrolled cell proliferation  
XX Disclosure; Table 2; 42pp; English.  
XX The patent claims a polypeptide isolated from the surface of Swiss 3T3  
CC cells which binds selectively with polypeptides of the bombesin type and  
CC binds with antagonist A and antagonist D. Antagonist A is a  
CC commercially available structural variant of substance P, known as  
CC [D-Arg1, D-Pro2, D-Trp7,9, Leu11] substance P. It is also known as



CC [D-Pro2] spantide. Antagonist B is also commercially available structural  
CC variant of substance P, known as [D-Phe5] spantide. Substance P is an  
CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten  
CC substance P antagonists (see AAP80313-80322) were tested for their  
CC ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue  
CC of bombesin in Swiss 3T3 cells). Antagonist D was clearly the most  
CC potent GRP antagonist. Peptides B, C, D, E, F, G, H, J and K were less  
CC potent than either A or D. Spantide (B) had no antagonist activity even  
CC at 100 uM. Polypeptide antagonists A and D and novel variants are useful  
CC for diagnosis and therapy, esp. of cancers where uncontrolled cell  
CC growth is associated with disorders of proteins of the bombesin family.  
XX Sequence 11 AA;

Query Match 67.6%; Score 48; DB 9; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RPKPQOWFWLM 11  
|||||: ||  
Db 1 rpkpqgffglm 11

RESULT 48  
AAP80320  
ID AAP80320 standard; protein; 11 AA.  
XX AAP80320;  
AC  
DT 14-SEP-1990 (first entry)  
XX  
XX Sequence of neuropeptide antagonist H which binds with polypeptide  
DE receptor for bombesin type polypeptides.  
DE  
XX  
KW Spantide; neuropeptide; polypeptide receptor; cancer diagnosis;  
KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides;  
KW antagonist H.  
XX  
XX Swiss 3T3 cells.  
OS

Key Location/Qualifiers  
FT Misc-difference 1 /label=OTHER  
FT /note="DArg"  
FT Misc-difference 2 /label=OTHER  
FT /note="DPro"  
FT Misc-difference 7 /label=OTHER  
FT /note="DPhe"  
FT Misc-difference 9 /label=OTHER  
FT /note="DHIS"  
FT Misc-difference 11 /label=OTHER  
FT /note="Met-NH2"

XX WO8807551-A.  
PN  
XX  
XX 06-OCT-1988.  
PD  
XX  
XX 31-MAR-1988; 88WO-GB00255.  
PF  
XX  
XX 25-NOV-1987; 87GB-0027638.  
PR  
XX  
XX (IMCR ) IMPERIAL CANCER RES.  
PA  
XX  
XX Rosengurt E, Zachary I, Woll P;  
PI  
XX  
XX WPI; 1988-292842/41.  
DR  
XX  
XX New polypeptide receptor for bombesin type polypeptide(s) -

PT is isolated from surface of Swiss 3T3 cells, and antibodies and  
PT antagonists are useful for treating uncontrolled cell proliferation  
XX  
XX Disclosure; Table 2; 42pp; English.  
XX  
CC The patent claims a polypeptide isolated from the surface of Swiss 3T3  
CC cells which binds selectively with polypeptides of the bombesin type and  
CC binds with antagonist A and antagonist D. Antagonist A is a  
CC commercially available structural variant of substance P, known as  
CC [D-Arg1, D-Pro2, D-Trp7, 9, Leu11] substance P. It is also known as  
CC [D-Pro2] spantide. Antagonist B is also commercially available structural  
CC variant of substance P, known as [D-Phe5] spantide. Substance P is an  
CC 11-mer neuropeptide, of interest in studies in pain transmission. Ten  
CC substance P antagonists (see AAP80313-80322) were tested for their  
CC ability to inhibit mitogenesis stimulated by GRP (the mammalian homologue  
CC of bombesin in Swiss 3T3 cells). Antagonist D was clearly the most  
CC potent GRP antagonist. Peptides B, C, D, E, F, G, H, J and K were less  
CC potent than either A or D. Spantide (B) had no antagonist activity even  
CC at 100 uM. Polypeptide antagonists A and D and novel variants are useful  
CC for diagnosis and therapy, esp. of cancers where uncontrolled cell  
CC growth is associated with disorders of proteins of the bombesin family.  
XX Sequence 11 AA;

Query Match 67.6%; Score 48; DB 9; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQOWFWLM 11  
|||||: ||  
Db 1 rpkpqgffglm 11

RESULT 49  
AAR13162  
ID AAR13162 standard; Protein; 11 AA.  
XX AAR13162;  
AC  
DT 10-OCT-1991 (first entry)  
XX  
XX Sialic acid-bonded polypeptide (2).  
DE  
XX Sialic acid; cataract; immune disorder.  
KW  
XX Synthetic.  
OS

Key Location/Qualifiers  
FT Modified-site 1 /note= "N-terminally glycosylated by 5-acetamido-  
FT 2,4,7,8,9-penta-O-acetyl-3,5-deoxy-beta-  
FT D-glycero-D-galactononulopyranosyl"  
XX JP03151398-A.  
PN  
XX 27-JUN-1991.  
PD  
XX 06-NOV-1989; 89JP-0288560.  
PF  
XX 06-NOV-1989; 89JP-0288560.  
PR  
XX (MECT-) MECT KK.  
PA  
XX WPI; 1991-233839/32.  
DR

XX New sialic acid derivs. bonded to physiologically active  
PT polypeptide - for treatment of cataracts, immune disorders etc.  
PT with prolonged half-life  
XX  
XX Example 4; Page 6; 7pp; Japanese.  
PS  
XX The prod. has prolonged half-life and is used as a pharmaceutical  
CC

CC for treatment of various diseases, such as cataract and immune  
CC disorders. It comprises a peptide, N-terminally glycosylated by  
CC (opt. acetylated) sialic acid.  
CC See also AAR12932, AAR13162 and AAR13201.

XX  
SQ Sequence 11 AA;  
  
Query Match 67.6%; Score 48; DB 12; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 1 rpkpqffglm 11

RESULT 50  
AAR11854  
ID AAR11854 standard; peptide; 11 AA.  
XX  
AC AAR11854;  
XX  
DT 09-JUL-1991 (first entry)  
XX  
DE Undecapeptide substance P.  
XX  
KW Undecapeptide; pharmaceutical; stress; sleep.  
XX  
OS Synthetic.  
XX  
PN DD285097-A.  
XX  
PD 05-DEC-1990.  
XX  
PF 21-JUN-1989; 89DD-0329831.  
XX  
PR 21-JUN-1989; 89DD-0329831.  
XX  
PA (DEAK ) INST WIRKSTOFF AKAD.  
XX (FARF ) VEB CHEM BITTERFELD.  
XX  
PI Beyermann M, Bienert M, Egler H, Haupke K, Krause E;  
XX Schwarz J, Walz H;  
XX WPI; 1991-133498/19.  
XX

PT Undeca-peptide substance pharmaceutical intermediate prepn. - by  
PT forming dipeptide between nitro-arginine and proline and  
PT reacting with polymer-bound non-peptide  
XX  
PS Calim 1; Page 1; 8pp; German.  
XX  
CC The peptide is prepared by solid phase synthesis.  
CC It can be used in the preparation of pharmaceuticals which can be  
CC used to treat certain stress-induced disturbances of the sleep  
CC profile.  
XX  
SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 12; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 1 rpkpqffglm 11

RESULT 51  
AAR21938  
ID AAR21938 standard; Protein; 11 AA.

XX  
AC AAR21938;  
XX  
DT 25-JUN-1992 (first entry)  
XX  
DE Substance P [Me-Leu 10].  
XX  
KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 10  
FT /label= OTHER  
FT /note= "OTHER = Me-Leu"  
XX  
PN W09202248-A.  
XX  
PD 20-FEB-1992.  
XX  
PF 29-JUL-1991; 91WO-US05323.  
XX  
PR 27-JUL-1990; 90US-0559173.  
XX  
PA (CHIL-) CHILDRENS MED CENT.  
XX  
PI Yankner BA;  
XX  
DR WPI; 1992-079804/10.  
XX  
PT Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX  
PS Claim 10; Page 21; 35pp; English.  
XX  
CC The peptide is the tachykinin agonist substance P with Me-Leu  
CC substituted at position 10. The peptide was synthesised  
CC by standard solid phase synthesis. Neuronal accumulation of  
CC beta-amyloid may be treated by administration of tachykinin  
CC agonists. The peptides can reduce the neurotoxic effects of a  
CC beta-amyloid related polypeptide on cultured neurons. The peptide  
CC and its analogues are useful for controlling diseases characterised  
CC by beta amyloid accumulation in the brain such as Alzheimer's  
CC disease and Down's syndrome.  
CC See also AAR21932-75.  
XX  
SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 13; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 1 rpkpqffglm 11

RESULT 52  
AAR21942  
ID AAR21942 standard; Protein; 11 AA.  
XX  
AC AAR21942;

XX  
DT 25-JUN-1992 (first entry)  
XX  
DE Substance P [MeMet 11].

XX  
KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX

```
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 11
XX /label= OTHER
XX /note= "OTHER = Methyl Methionine"
XX
XX WO9202248-A.
XX
XX 20-FEB-1992.
XX
XX 29-JUL-1991; 91WO-US05323.
XX
XX 27-JUL-1990; 90US-0559173.
XX
XX (CHIL-) CHILDRENS MED CENT.
XX
XX Yankner BA;
XX
XX WPI; 1992-079804/10.
XX
XX Treatment of neuronal accumulation of beta-amyloid - using
XX tachykinin agonists e.g. substance P, physalaemin and neurokinin
XX B, for treating Alzheimer's disease, Downs syndrome, etc.
XX
XX Claim 10; Page 21; 35pp; English.
XX
XX The peptide is the tachykinin agonist substance P with a methyl
XX methionine residue substituted at position 11. The peptide was
XX synthesised by standard solid phase synthesis. Neuronal
XX accumulation of beta-amyloid may be treated by administration of
XX tachykinin agonists. The peptide can reduce the neurotoxic effects
XX of a beta-amyloid related polypeptide on cultured neurons. The
XX peptide and its analogues are useful for controlling diseases
XX characterised by beta amyloid accumulation in the brain such as
XX Alzheimer's disease and Down's syndrome.
XX See also AAR21932-75.
XX
XX Sequence 11 AA;
XX
XX Query Match 67.6%; Score 48; DB 13; Length 11;
XX Best Local Similarity 81.8%; Pred. No. 0.25;
XX Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 RPKPOQWFWM 11
XX |||||:| |
XX Db 1 rpkpqgffglm 11
XX
XX RESULT 53
XX AAR21946
XX ID AAR21946 standard; Protein; 11 AA.
XX
XX AC AAR21946;
XX
XX DT 25-JUN-1992 (first entry)
XX
XX DE Substance P [Me-Phe 8].
XX
XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;
XX syndrome; hereditary cerebral haemorrhage.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 8
XX /label= OTHER
XX /note= "OTHER = Methyl phenylalanine"
XX
XX WO9202248-A.
XX
XX 20-FEB-1992.
XX
XX Query Match 67.6%; Score 48; DB 13; Length 11;
XX Best Local Similarity 81.8%; Pred. No. 0.25;
XX Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 RPKPOQWFWM 11
XX |||||:| |
XX Db 1 rpkpqgffglm 11
XX
XX RESULT 54
XX AAR21954
XX ID AAR21954 standard; Protein; 11 AA.
XX
XX AC AAR21954;
XX
XX DT 25-JUN-1992 (first entry)
XX
XX DE Substance P [Me-Gly 9].
XX
XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;
XX syndrome; hereditary cerebral haemorrhage.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 9
XX /label= OTHER
XX /note= "OTHER = Methyl glycine"
XX
XX WO9202248-A.
XX
XX 20-FEB-1992.
XX
XX 29-JUL-1991; 91WO-US05323.
XX
XX 27-JUL-1990; 90US-0559173.
XX
XX (CHIL-) CHILDRENS MED CENT.
XX
XX Yankner BA;
XX
XX WPI; 1992-079804/10.
XX
XX Treatment of neuronal accumulation of beta-amyloid - using
XX tachykinin agonists e.g. substance P, physalaemin and neurokinin
XX B, for treating Alzheimer's disease, Downs syndrome, etc.
XX
XX Claim 10; Page 21; 35pp; English.
XX
XX The peptide is the tachykinin agonist substance P with a methyl
XX phenylalanine residue substituted at position 8. The peptide was
XX synthesised by standard solid phase synthesis. Neuronal
XX accumulation of beta-amyloid may be treated by administration of
XX tachykinin agonists. The peptide can reduce the neurotoxic effects
XX of a beta-amyloid related polypeptide on cultured neurons. The
XX peptide and its analogues are useful for controlling diseases
XX characterised by beta amyloid accumulation in the brain such as
XX Alzheimer's disease and Down's syndrome.
XX See also AAR21932-75.
XX
XX Sequence 11 AA;
XX
XX Query Match 67.6%; Score 48; DB 13; Length 11;
XX Best Local Similarity 81.8%; Pred. No. 0.25;
XX Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 RPKPOQWFWM 11
XX |||||:| |
XX Db 1 rpkpqgffglm 11
XX
XX RESULT 54
XX AAR21954
XX ID AAR21954 standard; Protein; 11 AA.
XX
XX AC AAR21954;
XX
XX DT 25-JUN-1992 (first entry)
XX
XX DE Substance P [Me-Gly 9].
XX
XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;
XX syndrome; hereditary cerebral haemorrhage.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Misc-difference 9
XX /label= OTHER
XX /note= "OTHER = Methyl glycine"
XX
XX WO9202248-A.
XX
XX 20-FEB-1992.
XX
XX 29-JUL-1991; 91WO-US05323.
XX
XX 27-JUL-1990; 90US-0559173.
XX
XX (CHIL-) CHILDRENS MED CENT.
XX
XX Yankner BA;
XX
XX WPI; 1992-079804/10.
XX
```

XX Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX  
PS Claim 10; Page 22; 35pp; English.  
XX  
CC The peptide is the tachykinin agonist substance P with a methyl  
CC glycine residue substituted at position 9. The peptide was  
CC synthesised by standard solid phase synthesis. Neuronal  
CC accumulation of beta-amyloid may be treated by administration of  
CC tachykinin agonists. The peptide can reduce the neurotoxic effects  
CC of a beta-amyloid related polypeptide on cultured neurons. The  
CC peptide and its analogues are useful for controlling diseases  
CC characterised by beta amyloid accumulation in the brain such as  
CC Alzheimer's disease and Down's syndrome.  
CC See also AAR21932-75.  
XX  
SQ Sequence 11 AA;  
Query Match 67.6%; Score 48; DB 13; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
Db 1 rpkpqgffglm 11  
RESULT 55  
AAR21962  
ID AAR21962 standard; Peptide; 11 AA.  
XX  
AC AAR21962;  
XX  
DT 25-JUN-1992 (first entry)  
XX  
DE Substance P [Me Gly 6, Met (O2) 11].  
XX  
KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 6 /label= OTHER  
FT /note= "OTHER - Methyl glycine"  
FT  
FT Misc-difference 11  
FT /label= OTHER  
FT /note= "OTHER - Met (O2)"  
XX  
PN WO9202248-A.  
XX  
PD 20-FEB-1992.  
XX  
PF 29-JUL-1991; 91WO-US05323.  
XX  
PR 27-JUL-1990; 90US-0559173.  
XX  
PA (CHIL-) CHILDRENS MED CENT.  
XX  
PI Yankner BA;  
XX  
DR WPI; 1992-079804/10.  
XX  
PT Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX  
PS Claim 10; Page 22; 35pp; English.  
XX

CC The peptide is the tachykinin agonist, substance P with methyl  
CC glycine substituted at position 9 and Met (O2) at position 11.  
CC The peptide was synthesised by standard solid phase synthesis.  
CC Neuronal accumulation of beta-amyloid may be treated by administ-  
CC ration of tachykinin agonists. The peptide can reduce the neuro-  
CC toxic effects of a beta-amyloid related polypeptide on cultured  
CC neurons. The peptide and its analogues are useful for controlling  
CC diseases characterised by beta amyloid accumulation in the brain  
CC such as Alzheimer's disease and Down's syndrome.  
CC See also AAR21932-75.  
XX  
SQ Sequence 11 AA;  
Query Match 67.6%; Score 48; DB 13; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
Db 1 rpkpqgffglm 11  
RESULT 56  
AAR21963  
ID AAR21963 standard; Peptide; 11 AA.  
XX  
AC AAR21963;  
XX  
DT 25-JUN-1992 (first entry)  
XX  
DE Substance P [p-Chloro-Phe 7,8].  
XX  
KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 7 /label= OTHER  
FT /note= "OTHER - p-Chloro-phenylalanine"  
FT  
FT Modified-site 8 /label= OTHER  
FT /note= "OTHER - p-Chloro-phenylalanine"  
XX  
PN WO9202248-A.  
XX  
PD 20-FEB-1992.  
XX  
PF 29-JUL-1991; 91WO-US05323.  
XX  
PR 27-JUL-1990; 90US-0559173.  
XX  
PA (CHIL-) CHILDRENS MED CENT.  
XX  
PI Yankner BA;  
XX  
DR WPI; 1992-079804/10.  
XX  
PT Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX  
PS Claim 10; Page 22; 35pp; English.  
XX  
CC The peptide is the tachykinin agonist, substance P fragment  
CC with p-Chloro-phenylalanine residues substituted at positions 7 and  
CC 8. The peptide was synthesised by standard solid phase synthesis.  
CC Neuronal accumulation of beta-amyloid may be treated by administ-  
CC ration of tachykinin agonists. The peptide can reduce the neuro-  
CC toxic effects of a beta-amyloid related polypeptide on cultured  
CC neurons. The peptide and its analogues are useful for controlling

CC diseases characterised by beta amyloid accumulation in the brain  
 CC such as Alzheimer's disease and Down's syndrome.  
 CC See also AAR21932-75.

XX SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 13; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
 |||||:|  
 Db 1 rpkpqffglm 11

RESULT 57

AAR28442  
 ID AAR28442 standard; peptide; 11 AA.

XX AC AAR28442;

XX DT 22-MAR-1993 (first entry)

XX DE Substance P.

XX KW NK1 receptor; tumour; malignant glioma; pheochromocytoma;  
 KW paraganglia; small cell lung cancer; nerve regeneration; lymphoma;  
 KW granuloma; Crohn's disease.

XX OS Synthetic.

XX FH Key Location/Qualifiers  
 XX FT Modified-site 11 /note= "amidated"

XX PN WO9218536-A.

XX PD 29-OCT-1992.

XX PF 22-APR-1992; 92WO-US03307.

XX PR 22-APR-1991; 91EP-0200955.

XX PA (MLCW ) MALLINCKRODT MEDICAL INC.

XX PI Bakker WH, Hagen PM, Krenning EP, Lamberts SWJ, Visser TJ;

XX DR WPI; 1992-382047/46.

XX PT Detection and localisation of tissues with neurokinine-1 receptors  
 PT - for detecting and treating tumours having neurokinine-1  
 PT receptors e.g. malignant glioma, small cell lung cancer etc.

XX PS Disclosure; Page 4; 22pp; English.

XX CC Substance P or its Tyr0 deriv. is a preferred peptide having a  
 CC selective affinity to neurokinine-1 receptors which (when  
 CC labelled with a radioactive isotope) can be used in imaging methods.  
 CC A generic formula for preferred peptides is AAR28441. Such peptides  
 CC are thus useful in diagnosis and treatment of conditions that are  
 CC related to NK1 receptors and in visualising NK1 receptors on certain  
 CC tissues. See also AAR28443-R28446.

XX SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 13; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
 |||||:|

Db 1 rpkpqffglm 11

RESULT 58

AAR42646  
 ID AAR42646 standard; peptide; 11 AA.

XX AC AAR42646;

XX DT 19-APR-1994 (first entry)

XX DE Neurokinin 1 receptor affinity-contg. peptide (Substance P).

XX KW Neurokinin 1; somatostatin; receptor; cytokine; growth factor;  
 KW hormone; intra-operativ; tumour; low energy gamma photon;  
 KW radionuclide.

XX OS Synthetic.

XX FH Key Location/Qualifiers  
 XX FT Modified-site 11 /note= "the C-terminal is amidated"

XX PN WO9318797-A.

XX PD 30-SEP-1993.

XX PF 24-MAR-1993; 93WO-US02772.

XX PR 25-MAR-1992; 92EP-0200848.

XX PA (MLCW ) MALLINCKRODT MEDICAL INC.

XX PI Doedens BJ, Ensing GJ, Panek KJ;

XX DR WPI; 1993-320461/40.

XX PT Intra-operatively detecting and locating tumour tissues - using  
 PT specific peptide(s) labelled with low energy gamma photon  
 PT emitting radionuclide

XX PS Disclosure; Page 4; 31pp; English.

XX CC The method of intraoperatively detecting and locating tumoral  
 CC tissues makes use of peptides having selective neurokinin 1  
 CC receptor affinity (AAR42644: generic formula; AAR42646-R42650:  
 CC specific examples), peptides having selective somatostatin  
 CC receptor affinity (AAR42645: generic formula; AAR42651-R42660:  
 CC specific examples), and peptides selected from cytokines,  
 CC growth factors and hormones.

XX SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 14; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
 |||||:|  
 Db 1 rpkpqffglm 11

RESULT 59

AAR42647  
 ID AAR42647 standard; peptide; 11 AA.

XX AC AAR42647;

XX DT 19-APR-1994 (first entry)

XX DE Neurokinin 1 receptor affinity-contg. peptide.

KW Neurokinin 1; somatostatin; receptor; cytokine; growth factor;  
 KW hormone; intra-operativ; tumour; low energy gamma photon;  
 OS radionuclide.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 9

FT Modified-site 11

FT Modified-site 11 /label= MeGly

FT /note= "Met is Met(O2); the C-terminal is amidated"

XX W09318797-A.

XX 30-SEP-1993.

XX 24-MAR-1993; 93WO-US02772.

XX 25-MAR-1992; 92EP-0200848.

XX (MLCW ) MALLINCKRODT MEDICAL INC.

XX Doedens BJ, Ensing GJ, Panek KJ;

XX WPI; 1993-320461/40.

XX Intra-operatively detecting and locating tumour tissues - using

XX specific peptide(s) labelled with low energy gamma photon

XX emitting radionuclide

XX Disclosure; Page 5; 31pp; English.

XX The method of intraoperatively detecting and locating tumoral

XX tissues makes use of peptides having selective neurokinin 1

XX receptor affinity (AA42644; generic formula; AA42646-R42650;

XX specific examples), peptides having selective somatostatin

XX receptor affinity (AA42645; generic formula; AA42651-R42660;

XX specific examples), and peptides selected from cytokines,

XX growth factors and hormones.

XX Sequence 11 AA;

XX Query Match 67.6%; Score 48; DB 14; Length 11;

XX Best Local Similarity 81.8%; Pred. No. 0.25;

XX Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

XX QY 1 RPKPQOWFWLM 11

XX DB 1 rpkpqgffxlm 11

XX RESULT 60

XX AAR85243

XX ID AAR85243 standard; peptide; 11 AA.

XX AAR85243;

XX 18-AUG-1997 (first entry)

XX Substance P peptide.

XX Ligand; antibody; receptor; SELEX; random library; amplification; PCR;

XX Systematic Evolution of Ligands by EXponential enrichment; primer;

XX polymerase chain reaction; structure; mimicry; substance P; tachykinin;

XX neuropeptide; rheumatoid arthritis; atherosclerosis; cancer;

XX diabetic retinopathy.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 11

FT /note= "contains C-terminal NH2 group"

XX W09330775-A1.

XX 16-NOV-1995.

XX 03-MAY-1995; 95WO-US05600.

XX 21-DEC-1994; 94US-0361795.

XX 06-MAY-1994; 94US-0238863.

XX 24-MAY-1994; 94US-0248632.

XX 09-SEP-1994; 94US-0303362.

XX 11-JUN-1990; 90US-0536428.

XX 10-JUN-1991; 91US-0714131.

XX 21-OCT-1992; 92US-0964624.

XX (UYRE-) UNIV RES CORP.

XX Allen P, Doudna JA, Feigon J, Gold L, Nieuwlandt D;

XX Schneider DJ, Sullenger BA, Wecker M;

XX WPI; 1995-404132/51.

XX Systematic evolution of ligands by exponential enrichment - for

XX identifying nucleic acid ligands used in the treatment of, e.g. type

XX B insulin resistance and HIV

XX Example 10; Fig 8; 209pp; English.

XX The invention relates to a novel method of isolating ligands that bind

XX to target proteins e.g. antibodies or receptors, which bind other

XX proteins or ligands. The method, designated Systematic Evolution of

XX Ligands by EXponential enrichment (SELEX), comprises generating a library

XX of random oligonucleotide sequences, about 40-60 nucleotides in length,

XX and binding these sequences to the target proteins. After removal of

XX unbound material, the remaining bound nucleotides sequences are amplified

XX e.g. by PCR, and the newly amplified material is bound again with the

XX target protein. This cycle continues until a sufficiently pure

XX oligonucleotide sequence is isolated. The method allows the isolation of

XX oligonucleotide sequences which structurally mimic the target protein's

XX ligand. Ligands AAT06098-130 are examples of nucleic acid ligands which

XX bind the tachykinin-family neuropeptide Substance P (this sequence). The

XX new ligands were split into 2 groups based on their affinities for

XX Substance P. Class 1 ligands had binding affinities up to 2 micromolar

XX whereas class 2 ligands bound at above 2 micromolar. This sequence

XX represents the consensus of the class 1 ligands. The ligands can be

XX used to block the activity of Substance P and is useful in the treatment

XX of e.g. rheumatoid arthritis, atherosclerosis, diabetic retinopathy or

XX cancer.

XX Sequence 11 AA;

XX Query Match 67.6%; Score 48; DB 16; Length 11;

XX Best Local Similarity 81.8%; Pred. No. 0.25;

XX Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

XX QY 1 RPKPQOWFWLM 11

XX DB 1 rpkpqgffgml 11

XX RESULT 61

XX AAR77310

XX ID AAR77310 standard; peptide; 11 AA.

XX AAR77310;

XX 08-MAR-1996 (first entry)

XX Substance P.

XX Substance P; neurokinin; neurokinin receptor antagonist;

XX sensory perception; tachykinin receptor; therapy;

KW neurodegenerative disorder; Alzheimer's disease; demyelinating disease;  
KW multiple sclerosis; respiratory disease; ophthalmic disease;  
KW addiction disorder; adverse immune reaction; gastrointestinal disorder;  
KW bladder function disorder; fibrosing disease; collagen disease;  
KW diagnosis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 11  
FT Modified-site /note= "amided"  
XX  
XX  
PN US5434158-A.  
XX  
XX 18-JUL-1995.  
XX  
XX 26-APR-1994; 94US-0233487.  
XX  
XX 26-APR-1994; 94US-0233487.  
XX (MERI ) MERCK & CO INC.  
XX  
XX Shah SK;  
PI  
XX  
XX WPI; 1995-268290/35.  
XX  
XX New 1'-substd. spiro-indoline-3,4'-piperidine derivs. - useful as  
PT selective neurokinin-3 antagonists, e.g. for treating CNS disorders,  
PT migraine or esp. asthma.  
XX  
XX Disclosure; Column 1; 16pp; English.  
XX  
XX This sequence represents Substance P. This sequence, and those shown in  
CC AAR77311 and AAR77312 are tachykinins. These three sequences are  
CC pharmacologically active neuropeptides, and are neurokinin receptor  
CC agonists. Neurokinin receptors are widely distributed throughout the  
CC mammalian nervous system, circulatory system and peripheral tissues.  
CC Neurokinin receptors are involved in sensory perception. These  
CC sequences were used in the design and testing of neurokinin antagonists.  
CC These antagonists could be used in the treatment of conditions  
CC characterised by overstimulation of tachykinin receptors. The  
CC antagonists can also be used, for the treatment of neurodegenerative  
CC disorders (e.g. Alzheimer's disease), demyelinating diseases (e.g.  
CC multiple sclerosis), respiratory diseases, ophthalmic diseases, addiction  
CC disorders, adverse immune reactions, gastrointestinal disorders, bladder  
CC function disorders, fibrosing and collagen diseases. The antagonists can  
CC also be used as diagnostic agents.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 67.6%; Score 48; DB 16; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQOQFWLW 11  
Dd 1 rpkpqffglm 11

RESULT 62  
AAW33180  
ID AAW33180 standard; peptide; 11 AA.  
XX  
XX AAW33180;  
XX  
XX 29-JAN-1998 (first entry)  
XX  
XX Mono-DTPA-Arg1 Substance P.  
DE  
KW Substance P; radiolabel; diagnostic imaging; therapy;  
KW mono-DTPA-Arg1.  
XX

OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT Modified-site /note= "DTPA-Arg"  
FT Modified-site 11  
FT Modified-site /note= "amided"  
XX  
XX WO9640292-A1.  
XX  
XX 19-DEC-1996.  
XX  
XX 07-JUN-1996; 96WO-US09706.  
XX  
XX 07-JUN-1995; 95US-0480372.  
XX (MLCW ) MALLINCKRODT MEDICAL INC.  
XX Srinivasan A;  
XX WPI; 1997-087027/08.  
XX  
XX Prepn. of pure radio-labelled peptide, e.g. for diagnostic imaging -  
PT by combining protected poly-amino-carboxylate ligand with peptide  
PT and forming complex with radionuclide  
XX  
XX Example 3; Page 12; 20pp; English.  
XX  
XX Preparing a radiolabelled peptide composition, comprises combining  
CC a triamine or diamine chelating agent with a peptide, e.g. the  
CC present peptide, in a solid phase peptide synthesiser, and  
CC complexing a radionuclide with the chelate-peptide conjugate.  
CC Radiolabelled peptides or peptidomimetics can be used as diagnostic  
CC imaging agents, or in therapeutic applications, e.g. iodine(111)  
CC labelled pentatide can be used for somatostatin receptor  
CC imaging of neuroendocrine tumours. The radiolabelled products are  
CC obtained efficiently and inexpensively in high purity. The  
CC protected polyaminocarboxylate ligands can be added to the peptide  
CC by standard solution or solid phase peptide synthesis and  
CC deprotected with conventional reagents to give only the  
CC mono-addition product, free of di-addition product impurities. The  
CC deprotected product can be labelled with medically useful  
CC radionuclides, e.g. lanthanides or actinides, at any desired  
CC location. Pre-derivatisation of individual amino acids is not  
CC required.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 67.6%; Score 48; DB 18; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQOQFWLW 11  
Dd 1 rpkpqffglm 11

RESULT 63  
AAW04616  
ID AAW04616 standard; peptide; 11 AA.  
XX  
XX AAW04616;  
XX  
XX 13-AUG-1997 (first entry)  
XX  
XX Substance P peptide for mass spectrometry analysis.  
DE  
KW Mass spectrometry; polymer analysis; biopolymer analysis.  
KW Synthetic.  
XX  
XX WO9636986-A1.  
PN

XX 21-NOV-1996.  
 XX 17-MAY-1996; 96WO-US07146.  
 XX 19-MAY-1995; 95US-0447175.  
 XX 19-MAY-1995; 95US-0446055.  
 XX (PERS-) PERCEPTIVE BIOSYSTEMS INC.  
 XX Patterson DH, Tarr GE;  
 XX WPI; 1997-012308/01.  
 XX Sequencing polymers, e.g. DNA, RNA, peptide nucleic acids, proteins,  
 PT etc. - by obtaining mass to charge ratios of polymer fragments,  
 PT pref. using mass spectrometer, and performing statistical analysis  
 XX  
 PS Example 2; Page 32; 86pp; English.  
 XX A method of obtaining sequence information about a polymer (e.g. DNA,  
 CC RNA, peptide nucleic acids, proteins, peptides and carbohydrates)  
 CC comprising monomers of known mass has been claimed. The present  
 CC sequence represents a substance P peptide, and was used as  
 CC an example as a digestion before analysis by mass spectrometry,  
 CC using this novel on-plate strategy. Total sequence information  
 CC from a nine well digestion can be represented in a single digestion or  
 CC it is often derived from two or more wells. The methods, apparatus and  
 CC kit (claimed) can be used for the analysis of polymers, particularly  
 CC biopolymers, e.g. DNA, RNA, peptide nucleic acids, proteins, peptides  
 CC and carbohydrates. It provides a rapid, automated and cost effective  
 CC sequencing of polymers, with a statistical certainty.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 67.6%; Score 48; DB 18; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RPKPQQWFWM 11  
 Db 1 rpkpqgffglm 11  
 RESULT 64  
 AAW50978  
 ID AAW50978 standard; peptide; 11 AA.  
 XX AAW50978;  
 XX  
 DT 31-JUL-1998 (first entry)  
 DE Substance P analogue [D-Arg1,D-Pro2,D-Phe7,D-His9].  
 XX  
 KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
 KW Substance P; cancer; inhibition; growth hormone releasing factor;  
 KW spantide.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1 /note= "D-form residue"  
 FT Misc-difference 2 /note= "D-form residue"  
 FT Misc-difference 7 /note= "D-form residue"  
 FT Misc-difference 9 /note= "D-form residue"  
 FT Modified-site 11 /note= "D-form residue"  
 FT /note= "C-terminal amide"

PN EP835662-A2.  
 XX 15-APR-1998.  
 XX 11-DEC-1996; 96EP-0309012.  
 XX 08-OCT-1996; 96US-0727679.  
 XX 16-AUG-1996; 96IN-0001822.  
 XX (NAIM-) NAT INST IMMUNOLOGY.  
 XX Jaggi M, Mukherjee R;  
 XX WPI; 1998-208959/19.  
 XX Composition containing analogues of vasoactive intestinal peptide,  
 PT somatostatin - bombesin and substance P, for treatment of tumours  
 PT and for inhibiting over-expression of these peptide(s)  
 XX  
 PS Disclosure; Page 13; 49pp; English.  
 XX The invention relates to a new composition which comprises: (i) the  
 CC somatostatin analogue SOM2 AGCKNFRDQKPTSGC (3-14 disulphide bridge),  
 CC and (ii) at least 4 of the peptides: antagonist of vasoactive  
 CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP  
 CC antagonist (VIP3); somatostatin analogue (SOM1); bombesin  
 CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are  
 CC more general compositions containing peptide analogues of somatostatin,  
 CC VIP, bombesin and substance P. The compositions are used in human or  
 CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour  
 CC or cancer cells, particularly for treatment of leukaemia, lymphoma,  
 CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,  
 CC breast, kidney or particularly rectum and colon, and (b) to prevent,  
 CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
 CC cells express receptors for VIP, somatostatin, bombesin and/or substance  
 CC P. The present sequence represents a substance P analogue.  
 XX  
 SQ Sequence 11 AA;  
 Query Match 67.6%; Score 48; DB 19; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RPKPQQWFWM 11  
 Db 1 rpkpqgffhlm 11  
 RESULT 65  
 AAW42973  
 ID AAW42973 standard; Protein; 11 AA.  
 XX AAW42973;  
 XX  
 DT 01-MAY-1998 (first entry)  
 DE Substrate P reporter epitope.  
 XX  
 KW Beta-amyloid peptide; BAP; extracellular BAP plaque;  
 KW cerebrovascular deposit; Alzheimers disease; Downs syndrome;  
 KW amyloid precursor protein; APP; secretase; BAP aggregation;  
 KW abnormal proteolytic cleavage; substrate P reporter epitope.  
 XX  
 OS Synthetic.  
 XX  
 PN US5703209-A.  
 XX 30-DEC-1997.  
 XX  
 XX 05-JUN-1995; 95US-0464248.  
 XX 20-SEP-1993; 93US-0123659.





XX This sequence represents substance P used in the method of the  
 CC invention. The method is for enhancing opioid analgesia within a human  
 CC subject for a duration of 15 minutes comprises concurrent administration  
 CC of substance P, or one of its precursors. The method is used to elicit  
 CC opioid analgesia and anaesthesia, either prior to or after the occurrence  
 CC of a nociceptive event. The components have a synergistic effect. The  
 CC method allows use of low doses of opioid that produce little or no  
 CC physiological effect reducing conventional risks of toxicity and  
 CC addiction, and allows the use of low doses of substance P and its related  
 CC analogs that limit their in vivo physiological consequences. The  
 CC analgesia is naloxone reversible allowing diminishment or complete  
 CC elimination of opioid analgesia if desired and on demand. The treatment  
 CC provides a durable analgesic effect, but only minimally disturbs and  
 CC interrupts the normal metabolic processes of the body.

XX  
 SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 20; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
 |||||:| |  
 Db 1 rpkpqffglm 11

RESULT 68  
 AAW92715  
 ID AAW92715 standard; peptide; 11 AA.  
 XX  
 AC AAW92715;  
 DT 30-APR-1999 (first entry)  
 XX Human tachykinin agonist beta-amyloid peptide fragment #61.  
 DE Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
 KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
 KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
 XX  
 OS Homo sapiens.

Key Location/Qualifiers  
 Modified-site 10  
 /label= MeLeu  
 /note= "N-methyl-leucine"

US5876948-A.  
 02-MAR-1999.  
 27-JUL-1991; 91US-0737371.  
 29-JUL-1991; 91US-0737371.  
 27-JUL-1990; 90US-0559173.  
 (CHIL-) CHILDRENS MEDICAL CENT.  
 Yankner BA;  
 WPI; 1999-189630/16.

Screening for neurotoxin inhibitors - by testing compounds for their  
 effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
 Disclosure; Column 37-38; 28pp; English.

This invention describes a method for screening compounds for inhibiting  
 a neurotoxin. The method involves incubating tachykinin agonists with  
 neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
 used for identifying compounds for treating diseases characterised by an

CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
 CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
 CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
 CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
 CC beta-amyloid peptide fragments.

XX Sequence 11 AA;

Query Match 67.6%; Score 48; DB 20; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
 |||||:| |  
 Db 1 rpkpqffglm 11

RESULT 69  
 AAW92719  
 ID AAW92719 standard; peptide; 11 AA.  
 XX  
 AC AAW92719;  
 DT 30-APR-1999 (first entry)  
 XX Human tachykinin agonist beta-amyloid peptide fragment #65.  
 DE Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
 KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
 KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
 XX  
 OS Homo sapiens.

Key Location/Qualifiers  
 Modified-site 9  
 /label= MeGly  
 /note= "N-methyl-glycine"

US5876948-A.  
 02-MAR-1999.  
 27-JUL-1991; 91US-0737371.  
 29-JUL-1991; 91US-0737371.  
 27-JUL-1990; 90US-0559173.  
 (CHIL-) CHILDRENS MEDICAL CENT.  
 Yankner BA;  
 WPI; 1999-189630/16.

Screening for neurotoxin inhibitors - by testing compounds for their  
 effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
 Disclosure; Column 39-40; 28pp; English.

This invention describes a method for screening compounds for inhibiting  
 a neurotoxin. The method involves incubating tachykinin agonists with  
 neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
 used for identifying compounds for treating diseases characterised by an  
 undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
 Down's syndrome, and the syndromes of hereditary cerebral haemorrhage,  
 with amyloidosis and non-inherited congophilic angiopathy with cerebral  
 haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
 beta-amyloid peptide fragments.

XX Sequence 11 AA;

Query Match 67.6%; Score 48; DB 20; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 rpqpqffglm 11

## RESULT 70

AAW92720  
ID AAW92720 standard; peptide; 11 AA.

XX AC AAW92720;

XX AC 30-APR-1999 (first entry)

XX DE Human tachykinin agonist beta-amyloid peptide fragment #66.

XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 10  
FT /label= MeLeu  
FT /note= "N-methyl-leucine"

XX PN US5876948-A.

XX PD 02-MAR-1999.

XX PF 27-JUL-1991; 91US-0737371.

XX PR 29-JUL-1991; 91US-0737371.

XX PR 27-JUL-1990; 90US-0559173.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI Yankner BA;

XX DR WPI; 1999-189630/16.

XX PT Screening for neurotoxin inhibitors - by testing compounds for their  
FT effect on beta-amyloid peptide neurotoxic effect on neuronal cells

XX PS Disclosure; Column 39-40; 28pp; English.

XX CC This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.

XX SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 20; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 rpqpqffglm 11

## RESULT 71

AAW92708

ID AAW92708 standard; peptide; 11 AA.

XX AC AAW92708;

XX DT 30-APR-1999 (first entry)

XX DE Human tachykinin agonist beta-amyloid peptide fragment #54.

XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 7  
FT /note= "Modification results in p-chloro-Phe"  
FT Modified-site 8  
FT /note= "Modification results in p-chloro-Phe"

XX PN US5876948-A.

XX PD 02-MAR-1999.

XX PF 27-JUL-1991; 91US-0737371.

XX PR 29-JUL-1991; 91US-0737371.

XX PR 27-JUL-1990; 90US-0559173.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI Yankner BA;

XX DR WPI; 1999-189630/16.

XX PT Screening for neurotoxin inhibitors - by testing compounds for their  
FT effect on beta-amyloid peptide neurotoxic effect on neuronal cells

XX PS Disclosure; Column 33-34; 28pp; English.

XX CC This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.

XX SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 20; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 rpqpqffglm 11

## RESULT 72

AAW92680

ID AAW92680 standard; peptide; 11 AA.

XX AC AAW92680;

XX DT 30-APR-1999 (first entry)

XX DE Human tachykinin agonist beta-amyloid peptide fragment #26.

XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;



XX WPI; 1999-189630/16.  
XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX  
XX  
XX PS  
XX Disclosure; Column 19-20; 28pp; English.  
XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congenital angiodystrophy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 67.6%; Score 48; DB 20; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RPKPOQWFWM 11  
Db 1 rpkpqgffglm 11

RESULT 75  
AAW92731  
ID AAW92731 standard; peptide; 11 AA.  
XX  
AC AAW92731;  
XX  
DT 30-APR-1999 (first entry)  
XX  
XX Human tachykinin agonist beta-amyloid peptide fragment #77.  
XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congenital angiodystrophy.  
XX  
XX Homo sapiens.  
XX  
XX US5876948-A.  
PN  
XX  
XX 02-MAR-1999.  
PD  
XX  
XX 27-JUL-1991; 91US-0737371.  
PP  
XX  
XX 29-JUL-1991; 91US-0737371.  
PR  
XX 27-JUL-1990; 90US-0559173.  
PR  
XX (CHIL-) CHILDRENS MEDICAL CENT.  
PA  
XX  
XX Yankner BA;  
PI  
XX  
XX WPI; 1999-189630/16.  
DR  
XX  
XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX  
XX  
XX Disclosure; Column 43-44; 28pp; English.  
XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congenital angiodystrophy with cerebral

CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 67.6%; Score 48; DB 20; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RPKPOQWFWM 11  
Db 1 rpkpqgffglm 11

RESULT 76  
AAW9662  
ID AAW9662 standard; peptide; 11 AA.  
XX  
AC AAW9662;  
XX  
DT 02-MAR-1999 (first entry)  
XX  
XX Substance P derivative having complex glycosylation.  
DE  
XX Substance P; mannose; glycosylation; solubility.  
KW  
XX Synthetic.  
OS  
XX  
XX Key Location/Qualifiers  
FH Region 1..4  
FT /note= "optionally the first four residues may be  
FT deleted, leaving SP(5-11)"  
FT Modified-site 5 /note= "the side chain amide group is N-substituted  
FT with N-acetyl-D-glucosamine (GlcNAc) which in turn  
FT is extended in the 4-position with a complex type  
FT sugar chain, a high mannose type sugar chain or a  
FT mixed type sugar chain"  
FT Modified-site 11 /note= "Met-NH2, i.e. C-terminal amide"  
FT  
XX JP10306099-A.  
PN  
XX  
XX 17-NOV-1998.  
PD  
XX  
XX 28-NOV-1997; 97JP-0343979.  
PF  
XX  
XX 04-MAR-1997; 97JP-0065372.  
PR  
XX  
XX (NOGK ) 2H NOGUCHI KENKYUSHO.  
PA  
XX  
XX WPI; 1999-054306/05.  
DR  
XX  
XX New substance P derivatives with side chain containing sugar - has  
PT improved solubility  
PT  
XX  
XX Claim 1; Page 2; 8pp; Japanese.  
PS  
XX  
XX The sequence represents the peptide portion of a new Substance P  
CC derivative having complex glycosylation on the Gln(5) position. The  
CC derivative has improved solubility compared with Substance P.  
CC  
XX Sequence 11 AA;  
SQ

Query Match 67.6%; Score 48; DB 20; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RPKPOQWFWM 11  
Db 1 rpkpqgffglm 11

```

XX Key Location/Qualifiers
FH Modified-site 1
FT /note= "hydrogen attached"
FT Modified-site 11
FT /note= "amidated residue"
XX
PN W0200052147-A2.
XX
XX 08-SEP-2000.
XX
XX 03-MAR-2000; 2000WO-US05551.
XX
XX 05-MAR-1999; 99US-0123148.
XX
XX (UYGE-) UNIV GEORGIA RES FOUND INC.
XX (TRAV/) TRAVIS J.
XX (POTE/) POTEPA J.
XX (BANB/) BANBULA A.
XX
XX Travis J, Potempa J, Banbula A;
XX WPI; 2000-594181/56.
XX
XX Prolyl tripeptidyl-peptidase, active analog, fragment or variant useful
XX for identifying its inhibitor which is useful for protecting an animal
XX from a periodontal disease such as gingivitis and periodontitis -
XX
XX Example 4; Page 37; 58pp; English.
XX
XX The present sequence represents a substrate which was used to test
XX the activity of prolyl tripeptidyl-peptidases PTP-A and DPP IV. The
XX prolyl tripeptidyl-peptidase has an amidolytic activity, and cleaves
XX a peptide bond in a target polypeptide having at least 4 amino acids.
XX This bond is between a proline and an amino acid attached to the
XX alpha-carboxyl group end of the proline. The polypeptide is useful for
XX identifying inhibitors. These inhibitors are then useful for reducing
XX the growth of bacterium or for protecting an animal from a periodontal
XX disease such as gingivitis and periodontitis caused by Porphyromonas
XX gingivalis.
XX
XX Sequence 11 AA;
XX
XX Query Match 67.6%; Score 48; DB 21; Length 11;
XX Best Local Similarity 81.8%; Pred. No. 0.25;
XX Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
XX Qy 1 RPKPQQWFWM 11
XX |||||:|
XX Db 1 rpkpqqffglm 11
XX
XX RESULT 79
XX AAB23027
XX ID AAB23027 standard; peptide; 11 AA.
XX
XX AC AAB23027;
XX
XX 16-JAN-2001 (first entry)
XX
XX Human/rat tachykinin Substance P.
XX
XX Substance P; tachykinin; human; rat; magnesium binding defect;
XX sodium sensitive essential hypertension; insulin resistance;
XX type 2 diabetes; antibody; immunoassay; quantification.
XX
XX Homo sapiens.
XX OS Rattus sp.
XX
XX Key Location/Qualifiers
XX Modified-site 11
XX /note= "C-terminal amide"
XX

```

XX PN WO200054053-A1.  
 XX PD 14-SEP-2000.  
 XX PF 09-MAR-2000; 2000WO-US03707.  
 XX PR 10-MAR-1999; 99US-0265690.  
 XX PA (WELLS) WELLS I C.  
 XX PI Wells IC;  
 XX WPI; 2000-587457/55.  
 XX PT Detecting magnesium binding defects associated with abnormal  
 CC physiological states such as sodium-sensitive essential hypertension  
 CC and type 2 insulin-resistant diabetes mellitus, comprises measuring a  
 CC specific pentapeptide in blood -  
 XX Disclosure; Page 5; 21pp; English.  
 XX The invention relates to a method for detecting magnesium binding  
 CC defects. The method comprises quantitating a tachykinin C-terminal  
 CC pentapeptide (e.g., AAB23025) and its degradation products (e.g.,  
 CC AAB23026) in blood using an antibody specific for the generalised  
 CC mammalian tachykinin C-terminal pentapeptide  
 CC Phe-(Phe/Val)-Gly-Leu-Met-NH<sub>2</sub> (AAB23028). The method is useful for  
 CC detecting cellular magnesium binding defects which are associated with  
 CC abnormal physiological states such as sodium-sensitive essential  
 CC hypertension and type 2 diabetes mellitus. The present sequence  
 CC represents the tachykinin Substance P from human and rat. C-terminal  
 CC fragments (AAB23025, AAB23026) of the present sequence may be assayed  
 CC according to the method of the invention.  
 XX SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 21; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 RPKPQQWFWM 11  
 Db 1 rpkpqgfglm 11  
 RESULT 80  
 AAY32382  
 ID AAY32382 standard; Peptide; 11 AA.  
 XX AC AAY32382;  
 XX DT 28-FEB-2000 (first entry)  
 XX DE Cell differentiation, proliferation and maintenance factor peptide.  
 XX KW Cell differentiation; cell proliferation; cell maintenance;  
 KW ectoderm-like cell; embryonic stem cell; pluripotent cell;  
 KW gene therapy; cell therapy; tissue transplant; organ transplant;  
 KW xerotransplant; allotransplant; concomitant transplantation;  
 KW transgenic animal; substance P.  
 XX OS Synthetic.  
 XX PN WO9953021-A1.  
 XX PD 21-OCT-1999.  
 XX PF 09-APR-1999; 99WO-AU00265.  
 XX PR 09-APR-1998; 98AU-0002912.  
 PR 23-SEP-1998; 98AU-0006097.

XX PA (BRES-) BRESAGEN LTD.  
 XX PI Bettess MD, Rathjen PD, Rathjen J;  
 XX DR WPI; 2000-061970/05.  
 XX PT New isolated biologically active factor capable of influencing  
 CC differentiation, proliferation or maintenance of pluripotent cells  
 XX Claim 3; Page 123; 189pp; English.  
 XX This sequence represents a peptide (substance P free acid) that can  
 CC form the low mol.wt. component of a novel biologically active factor  
 CC that is capable of influencing the differentiation, proliferation  
 CC and/or maintenance of pluripotent cells. The factor consists of a  
 CC low mol.wt. component selected from Pro, Pro-Ala, Ala-Pro-Gly,  
 CC Pro-OH-Pro, Gly-Pro-Ala, Gly-Pro-OH-Pro, a peptide given in  
 CC AAY32378-82, or a protease digested (including collagenase digested)  
 CC collagen fragment, and a high mol.wt. component such as fibronectin.  
 CC The biologically active factor is obtained from conditioned media of  
 CC hepatic or hepatoma cells or cell lines or extraembryonic endodermal  
 CC cells or cell lines. The factor is capable of causing the  
 CC transition of pluripotent cells (e.g. embryonic stem cells in  
 CC adherent culture and in suspension culture) to pluripotent cells  
 CC having different properties, more specifically primitive  
 CC ectoderm-like (EPL) cells. The factor is also capable of  
 CC maintaining and supporting proliferation of these cells in vitro.  
 CC It also allows the isolation and maintenance of EPL cells derived  
 CC from in vitro and in vivo primitive ectoderm. These cells can be  
 CC used in allo-, concomitant- or xeno-transplantation, cell therapy,  
 CC tissue and organ augmentation or replacement, and gene therapy.  
 CC They can also be used for producing chimeric or transgenic animals.  
 XX SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 21; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.25;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 Qy 1 RPKPQQWFWM 11  
 Db 1 rpkpqgfglm 11  
 RESULT 81  
 AAG62768  
 ID AAG62768 standard; peptide; 11 AA.  
 XX AC AAG62768;  
 XX DT 17-SEP-2001 (first entry)  
 XX DE Amino acid sequence of substance P.  
 XX KW Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
 XX OS Unidentified.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 11  
 FT /note= "amidated residue"  
 XX PN WO200153336-A1.  
 XX PD 26-JUL-2001.  
 XX PF 17-JAN-2001; 2001WO-US01529.  
 XX PR 19-JAN-2000; 2000US-0489667.  
 XX PA (ALLR ) ALLERGAN SALES INC.

XX  
PI Donovan S;  
XX  
DR WPI; 2001-451900/48.  
XX  
PT Agent useful for treating pain comprises a clostridial neurotoxin (or  
PT component) attached to a targeting moiety -  
XX  
PS Disclosure; Page 61; 77pp; English.  
XX  
CC The specification describes an agent, comprising a clostridial neurotoxin  
CC attached to a targeting moiety, where the targeting moiety is selected  
CC from transmission compounds, and compounds substantially similar to the  
CC transmission compounds. The agent may be used for treating pain, where  
CC the clostridial neurotoxin component is derived from botulinum toxin  
CC selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
CC The targeting moiety comprises a light chain and an amine end segment of  
CC a heavy chain and comprises Substance P as the targeting moiety. The pain  
CC alleviating effects persist for 2-6 months. The present sequence  
CC represents substance P, and is used in the course of the invention.  
XX  
SQ Sequence 11 AA;  
  
Query Match 67.6%; Score 48; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPKPQOWFWLM 11  
DB 1 rpkpqffgflm 11  
  
RESULT 82  
AAG99354  
ID AAG99354 standard; Peptide; 11 AA.  
AC AAG99354;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
DE Substance P peptide.  
XX  
KW Atypical tachykinin; ATT; human; hypertension..  
XX  
OS Unidentified.  
XX  
PN WO200146415-A1.  
XX  
PD 28-JUN-2001.  
XX  
PF 21-DEC-2000; 2000WO-JP09083.  
XX  
PR 21-DEC-1999; 99JP-0362638.  
PR 10-MAR-2000; 2000JP-0066714.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Itoh Y, Nishi K, Kitada C, Inatomi N;  
XX  
DR WPI; 2001-441676/47.  
XX  
PT Atypical tachykinin peptides of human origin and DNA encoding them for  
PT screening potential agents for treatment of hypertension -  
XX  
PS Disclosure; Page 9; 153pp; Japanese.  
XX  
CC The present invention relates to atypical tachykinin proteins of human  
CC origin and their esters, amides, salts and partial peptides. These can be  
CC used in the treatment, prevention and diagnosis of hypertension. The  
CC present sequence is a protein fragment described in the exemplification  
CC of the invention.

SQ Sequence 11 AA;  
  
Query Match 67.6%; Score 48; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPKPQOWFWLM 11  
DB 1 rpkpqffgflm 11  
  
RESULT 83  
AAB84527  
ID AAB84527 standard; peptide; 11 AA.  
XX  
AC AAB84527;  
XX  
DT 05-SEP-2001 (first entry)  
XX  
DE Amino acid sequence of human substance P.  
XX  
KW Substance P; cell toxin; Pseudomonas exotoxin; cell ablation;  
KW NK-1 receptor; chronic pain; tumour; neurological dysfunction;  
KW basal ganglia; cholinergic interneuron; Parkinson's disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200131020-A1.  
XX  
PD 03-MAY-2001.  
XX  
PF 20-OCT-2000; 2000WO-US29064.  
XX  
PR 22-OCT-1999; 99US-0161159.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Fitzgerald DJ, Iadarola MJ;  
XX  
DR WPI; 2001-417560/44.  
XX  
PT Making cell toxin to treat chronic pain, by forming substance  
PT P-Pseudomonas exotoxin disulfide-linked conjugate, by reacting modified  
PT exotoxin and substance P having additional cysteine residue at its  
PT N-terminus -  
XX  
PS Disclosure; Page 10; 54pp; English.  
XX  
CC The present sequence represents a human substance P. The peptide is  
CC used to produce a cell toxin. The cell toxin comprises a substance  
CC P-Pseudomonas exotoxin disulfide-linked conjugate. The cell toxin is  
CC useful for ablating NK-1 receptor expressing cells, such as dorsal horn  
CC cell, a stratum cell or a brain parenchyma cell, for treating chronic  
CC pain in epineurium cells, perineurium cells, nerve ganglia, nerve  
CC sheaths, nerve linings, meninges, pia mater cells, arachnoid membrane  
CC cells, duramembrane cells, cells lining a joint or brain or spinal cord  
CC parenchymal cells, without significantly affecting basal nociceptive  
CC responses. The cell toxin is thus useful for treating chronic pain or  
CC tumours that binds substance P. It is also useful for neurological  
CC dysfunctions of the basal ganglia by targeting cholinergic interneurons  
CC that express substance P e.g. Parkinson's disease.  
XX  
SQ Sequence 11 AA;  
  
Query Match 67.6%; Score 48; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPKPQOWFWLM 11  
DB 1 rpkpqffgflm 11



```
RESULT 84
AAB98866
ID AAB98866 standard; Peptide; 11 AA.
XX
AC AAB98866;
XX
DT 14-AUG-2001 (first entry)
XX
DE Chimeric analgesic peptide #22.
XX
KW Opioid receptor binding; nociceptive receptor binding; analgesic;
KW pain; chimeric peptide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 11
FT /label= OTHER
FT /note= "C-terminal amide"
XX
PN WO200130371-A2.
XX
PD 03-MAY-2001.
XX
PF 27-OCT-2000; 2000WO-US29789.
XX
PR 28-OCT-1999; 99US-0428692.
XX
PA (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.
XX
PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;
XX
DR WPI; 2001-397593/42.
XX
XX New chimeric peptides used for treating pain comprise opioid receptor
binding group and nociceptive receptor binding group -
XX
PS Claim 10; Page 15; 34pp; English.
XX
CC The present invention describes a number of chimeric peptides comprising
an opioid receptor binding moiety and a nociceptive receptor binding
moieties. These can be used as analgesics for the treatment of pain. Unlike
opioid receptor based peptides alone, tolerance does not result from
their long-term use. The present sequence is one of the peptides of the
invention.
XX
SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 22; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.25;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWFWM 11
Db 1 rpkpqffglm 11

RESULT 85
AAB82070
ID AAB82070 standard; peptide; 11 AA.
XX
AC AAB82070;
XX
DT 22-JUN-2001 (first entry)
XX
DE Substance P.
XX
KW Antigen; immunostimulant; vaccine; pharmaceutical composition; antiviral;
KW viral infection; substance P.
XX
```

```
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Modified-site 11
FT /note= "C-terminal amide"
XX
PN WO200124822-A2.
XX
PD 12-APR-2001.
XX
PF 02-OCT-2000; 2000WO-EP09657.
XX
PR 01-OCT-1999; 99AT-0001680.
XX
PA (CIST-) CISTEM BIOTECHNOLOGIES GMBH.
XX
PI Fleitmann J, Mattner F, Buschle M, Melling J;
XX
DR WPI; 2001-290577/30.
XX
XX New pharmaceutical composition comprising an antigen, an
immunostimulating substance and a polycationic polymer, useful in
manufacturing vaccines -
XX
XX Example 3; Page 14; 20pp; English.
XX
CC The present invention relates to a pharmaceutical composition comprising
(a) an antigen; (b) an immunostimulating substance consisting of
neuroactive compounds, hormones, compounds having growth hormone activity
or their mixtures; and (c) a polycationic polymer. The composition is
useful in manufacturing vaccines. To illustrate the present invention, a
murine tyrosinase related protein-2 peptide (TRP-2 peptide; see
AAB82064), was used. Mice were injected subcutaneously with either the
TRP-2 peptide, TRP-2 peptide + poly-L-arginine 60 (pR60) or TRP-2 peptide
+ pR60 + substance P (the present peptide). Animals were sacrificed 10
days post injection, and spleen tissue was harvested. Lymphocytes were
prepared from the spleen tissue and were re-stimulated with TRP-2 peptide
or with an ovalbumin-derived peptide (AAB82065), with the same major
histocompatibility complex (MHC) restriction serving as negative control.
Spots representing single T cells specific for the peptide used for
re-stimulation were counted. No spots were detected when the ovalbumin
derived peptide was used, while TRP-2 peptide + pR60 + substance P showed
the highest number of spots or single T cells.
XX
SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 22; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.25;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWFWM 11
Db 1 rpkpqffglm 11

RESULT 86
AAB91411
ID AAB91411 standard; Peptide; 11 AA.
XX
AC AAB91411;
XX
DT 22-JUN-2001 (first entry)
XX
DE Tachykinins peptide SEQ ID NO:587.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidy1; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
```

PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
XF 17-MAY-2000; 2000WO-US13576.  
XX  
PF 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX  
PI WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
PS Disclosure; Page 392; 733pp; English.  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFHLM 11  
|||||:|  
Db 1 rpkpqqffhlm 11

RESULT 87  
AAB91436  
ID AAB91436 standard; Peptide; 11 AA.  
XX  
AC AAB91436;  
XX  
DT 22-JUN-2001 (first entry)  
DE  
DE Tachykinins peptide SEQ ID NO:612.  
XX  
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX

PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX  
PI WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
PS Disclosure; Page 399; 733pp; English.  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 11 AA;

Query Match 67.6%; Score 48; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFHLM 11  
|||||:|  
Db 1 rpkpqqffhlm 11

RESULT 88  
AAB91449  
ID AAB91449 standard; Peptide; 11 AA.  
XX  
AC AAB91449;  
XX  
DT 22-JUN-2001 (first entry)  
DE  
DE Tachykinins peptide SEQ ID NO:625.  
XX  
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.

PR 15-OCT-1999; 99US-0159783.  
XX (CONJ-) CONJUCHEM INC.  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX WPI; 2001-112059/12.  
XX Modifying and attaching therapeutic peptides to albumin prevents  
XX peptidase degradation, useful for increasing length of in vivo activity  
XX  
XX Disclosure; Page 403; 733pp; English.  
XX The present invention describes a modified therapeutic peptide (I)  
XX comprising a therapeutically active amino acid region (III) and a  
XX reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
XX bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX factors and neurotransmitters, to protect them from peptidase activity  
XX in vivo for the treatment of various disorders. Endogenous therapeutic  
XX peptides are not suitable as drug candidates as they require frequent  
XX administration due to rapid degradation by peptidases in the body.  
XX Modifying and attaching therapeutic peptides to albumin prevents or  
XX reduces the action of peptidases to increase length of activity (half  
XX life) and specificity as bonding to large molecules decreases  
XX intracellular uptake and interference with physiological processes.  
XX AAB90829 to AAB92441 represent peptides which can be used in the  
XX exemplification of the present invention.  
XX Sequence 11 AA;  
SQ

Query Match 67.6%; Score 48; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
DB 1 rpkpqgfglm 11

RESULT 89  
AAB91450  
ID AAB91450 standard; Peptide; 11 AA.  
XX  
XX AAB91450;  
DT 22-JUN-2001 (first entry)  
XX  
DE Tachykinins peptide SEQ ID NO:626.  
XX  
XX protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
XX WO200069900-A2.  
PN  
XX  
PD 23-NOV-2000.  
XX  
XX 17-MAY-2000; 2000WO-US13576.  
PF  
XX  
XX 17-MAY-1999; 99US-0134406.  
PR  
XX 10-SEP-1999; 99US-0153406.  
PR  
XX 15-OCT-1999; 99US-0159783.  
XX  
XX (CONJ-) CONJUCHEM INC.  
XX

PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX WPI; 2001-112059/12.  
XX Modifying and attaching therapeutic peptides to albumin prevents  
XX peptidase degradation, useful for increasing length of in vivo activity  
XX  
XX Disclosure; Page 403; 733pp; English.  
XX The present invention describes a modified therapeutic peptide (I)  
XX comprising a therapeutically active amino acid region (III) and a  
XX reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
XX bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX factors and neurotransmitters, to protect them from peptidase activity  
XX in vivo for the treatment of various disorders. Endogenous therapeutic  
XX peptides are not suitable as drug candidates as they require frequent  
XX administration due to rapid degradation by peptidases in the body.  
XX Modifying and attaching therapeutic peptides to albumin prevents or  
XX reduces the action of peptidases to increase length of activity (half  
XX life) and specificity as bonding to large molecules decreases  
XX intracellular uptake and interference with physiological processes.  
XX AAB90829 to AAB92441 represent peptides which can be used in the  
XX exemplification of the present invention.  
XX Sequence 11 AA;  
SQ

Query Match 67.6%; Score 48; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
DB 1 rpkpqgfglm 11

RESULT 90  
AAB50544  
ID AAB50544 standard; peptide; 11 AA.  
XX  
XX AAB50544;  
AC  
XX 16-MAR-2001 (first entry)  
DT  
XX  
DE Prolyl endopeptidase inhibitor substance P peptide.  
XX  
XX Prolyl endopeptidase inhibitor; PEP inhibitor; central nervous system;  
KW CNS; nootropic; brain function disorder; Alzheimer's disease; amnesia.  
XX  
XX Unidentified.  
XX  
XX Key Location/Qualifiers  
FT Modified-site 11 /note= "amidated"  
FT  
XX  
XX WO200071144-A1.  
PN  
XX  
XX 30-NOV-2000.  
PD  
XX  
XX 16-MAY-2000; 2000WO-JP03135.  
PF  
XX  
XX 19-MAY-1999; 99JP-0138791.  
PR  
XX  
XX (DOME-) DOME INC.  
PA  
XX  
XX Kayahara H, Tsukahara K, Inagaki T;  
PI  
XX  
XX WPI; 2001-070833/08.  
DR  
XX

PT Prolyl endopeptidase inhibitor comprises cereal extract including new  
XX ketone compound.  
XX  
XX Disclosure; Fig 1; 27pp; Japanese.  
XX  
XX The present invention describes prolyl endopeptidase (PEP) inhibitors  
CC comprising a cereal extract. Also described are:  
CC (i) a 7-octadecenyl-7,10-henecosadienyl ketone;  
CC (ii) germinating brown rice having prolyl endopeptidase inhibitory  
CC activity for preventing and/or relieving brain function disorders; and  
CC (iii) foods for preventing or relieving brain function disorders;  
CC comprising the above PEP inhibitor or the above germinated brown rice.  
CC The PEP inhibitors can have central nervous system (CNS) and nootropic  
CC activity. The PEP inhibitors can be used for preventing and relieving  
CC brain function disorders including Alzheimer's disease and amnesia.  
CC The present sequence represents a PEP inhibitor peptide given in the  
XX exemplification of the present invention.  
XX  
XX Sequence 11 AA;  
SQ  
Query Match 67.6%; Score 48; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
DB 1 rpkpqffglm 11  
RESULT 91  
AAB50306  
ID AAB50306 standard; peptide; 11 AA.  
XX  
AC AAB50306;  
XX  
DT 08-MAR-2001 (first entry)  
XX  
DE Substance P.  
XX  
KW Antibacterial; Botulinum toxin inhibitor; BttxB;  
KW Previn; tetanus neurotoxin; buforinin; substance P.  
XX  
OS Unidentified.  
XX  
XX WO2000069891-A2.  
XX  
XX 23-NOV-2000.  
XX  
XX 15-MAY-2000; 2000WO-US13215.  
XX  
XX 17-MAY-1999; 99US-0134446.  
XX  
XX (USSA ) US DEPT OF THE ARMY.  
XX  
XX Gordon RK, Moorad DR, Doctor BP, Garcia GE;  
XX  
XX WPI; 2001-025001/03.  
XX  
XX Novel Previn compounds useful for inhibiting the protease activity of  
PT Botulinum B and tetanus toxins -  
XX  
XX Claim 7; Page 29; 47pp; English.  
XX  
XX The present sequence was investigated in the search for Botulinum  
CC toxin inhibitors (BttxB). Previn compounds which inhibit the enzymatic  
CC activity of BttxB and tetanus neurotoxins were isolated. Previn  
CC may be used to construct compounds such as buforinins.  
XX  
XX Sequence 11 AA;  
SQ  
Query Match 67.6%; Score 48; DB 22; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.25;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
DB 1 rpkpqffglm 11  
RESULT 92  
AAR32798  
ID AAR32798 standard; peptide; 12 AA.  
XX  
AC AAR32798;  
XX  
DT 17-JUN-1993 (first entry)  
XX  
DE Tyr-1 substance P used for binding assay.  
XX  
KW human substance P receptor protein; SP; neurotransmitter;  
KW neuromodulator; central nervous system; peripheral nervous system;  
KW gastrointestinal disorders; inflammation; immune disease.  
XX  
OS Homo sapiens.  
XX  
XX WO9303137-A.  
XX  
XX 18-FEB-1993.  
XX  
XX 05-AUG-1992; 92WO-US06532.  
XX  
XX 07-AUG-1991; 91US-0741200.  
XX  
XX (UNIW ) UNIV WASHINGTON.  
XX  
XX Krause JE;  
XX  
XX WPI; 1993-076495/09.  
XX  
XX New human substance P receptor protein and DNA encoding it - used  
PT e.g. for screening substance P antagonists  
XX  
XX Example; Page 8; 40pp; English.  
XX  
XX This sequence represents Tyr-1 substance P and was used in its  
CC 125-Iodinated form in a ligand binding assay of COS-7 cells  
CC transfected with substance P receptor coding plasmids (see AAQ37210).  
XX  
XX Sequence 12 AA;  
SQ  
Query Match 67.6%; Score 48; DB 14; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.28;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
DB 2 rpkpqffglm 12  
RESULT 93  
AAR85244  
ID AAR85244 standard; peptide; 12 AA.  
XX  
XX AAR85244;  
XX  
XX 18-AUG-1997 (first entry)  
XX  
XX Substance P analogue peptide Cys-SP.  
XX  
XX Ligand; antibody; receptor; SELEX; random library; amplification; PCR;  
KW Systematic Evolution of Ligands by Exponential enrichment; primer;  
KW polymerase chain reaction; structure; mimicry; substance P; tachykinin;  
KW neuropeptide; rheumatoid arthritis; atherosclerosis; cancer;

diabetic retinopathy.  
Synthetic.  
Key Location/Qualifiers  
Modified-site 1  
/note= "Ac-Arg"  
XX WO9530775-A1.  
XX 16-NOV-1995.  
XX 03-MAY-1995; 95WO-US05600.  
XX 21-DEC-1994; 94US-0361795.  
XX 06-MAY-1994; 94US-0238863.  
XX 24-MAY-1994; 94US-0248632.  
XX 09-SEP-1994; 94US-0303362.  
XX 11-JUN-1990; 90US-0536428.  
XX 10-JUN-1991; 91US-0714131.  
XX 21-OCT-1992; 92US-0964624.  
XX (UYRE-) UNIV RES CORP.  
XX Allen P, Doudna JA, Feigon J, Gold L, Nieuwlandt D;  
XX Schneider DJ, Sullenger BA, Wecker M;  
XX WPI; 1995-404132/51.  
XX Systematic evolution of ligands by exponential enrichment - for  
XX identifying nucleic acid ligands used in the treatment of, e.g. type  
XX B insulin resistance and HIV  
XX Example 11; Fig 8; 209pp; English.  
XX The invention relates to a novel method of isolating ligands that bind  
XX to target proteins e.g. antibodies or receptors, which bind other  
XX proteins or ligands. The method, designated Systematic Evolution of  
XX Ligands by Exponential enrichment (SELEX), comprises generating a library  
XX of random oligonucleotide sequences, about 40-60 nucleotides in length,  
XX and binding these sequences to the target proteins. After removal of  
XX unbound material, the remaining bound nucleotide sequences are amplified  
XX e.g. by PCR, and the newly amplified material is bound again with the  
XX target protein. This cycle continues until a sufficiently pure  
XX oligonucleotide sequence is isolated. The method allows the isolation of  
XX oligonucleotide sequences which structurally mimic the target protein's  
XX ligand. This peptide represents an analogue of Substance P (AAR85243) in  
XX which the N-terminal amine has been acylated in order to determine  
XX whether this functional group interacts with nucleic acid ligands binding  
XX substance P (see AAT06098-130). The ligands can be used to block the  
XX activity of Substance P and is useful in the treatment of e.g. rheumatoid  
XX arthritis, atherosclerosis, diabetic retinopathy or cancer.  
XX  
XX Sequence 12 AA;  
Query Match 67.6%; Score 48; DB 16; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.28;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 RPKPQQWFWM 11  
Db 1 rpkpqffglm 11  
RESULT 94  
AAY03157  
ID AAY03157 standard; peptide; 12 AA.  
XX AAY03157;  
XX 10-JUN-1999 (first entry)  
XX  
Substance P-Glycine.  
Opioid peptide; opioid analgesia; enhancer; opioid anaesthesia;  
substance P.  
Synthetic.  
XX US5891842-A.  
XX 06-APR-1999.  
XX 12-APR-1996; 96US-0631434.  
XX 09-APR-1993; 93US-0044954.  
XX 12-APR-1996; 96US-0631434.  
XX (TUFT ) TUFTS COLLEGE.  
XX Kream RM;  
XX WPI; 1999-253906/21.  
XX Synergistic method for enhancing opioid analgesia and anaesthesia  
XX within a human  
XX Disclosure; Column 14; 20pp; English.  
XX This sequence represents substance P used in the method of the  
XX invention. The method is for enhancing opioid analgesia within a human  
XX subject for a duration of 15 minutes comprises concurrent administration  
XX of substance P, or one of its precursors. The method is used to elicit  
XX opioid analgesia and anaesthesia, either prior to or after the occurrence  
XX of a nociceptive event. The components have a synergistic effect. The  
XX method allows use of low doses of opioid that produce little or no  
XX physiological effect reducing conventional risks of toxicity and  
XX addiction, and allows the use of low doses of substance P and its related  
XX analogs that limit their in vivo physiological consequences. The  
XX analgesia is naloxone reversible allowing diminishment or complete  
XX elimination of opioid analgesia if desired and on demand. The treatment  
XX provides a durable analgesic effect, but only minimally disturbs and  
XX interrupts the normal metabolic processes of the body.  
XX  
XX Sequence 12 AA;  
Query Match 67.6%; Score 48; DB 20; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.28;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Oy 1 RPKPQQWFWM 11  
Db 1 rpkpqffglm 11  
RESULT 95  
AAW94412  
ID AAW94412 standard; peptide; 12 AA.  
XX AAW94412;  
XX 15-APR-1999 (first entry)  
XX Cancer protease-sensitive amino acid linker PAP-215 and PAP-216.  
XX Ricin-like toxin; cancer; viral infection; parasitic infection;  
XX linker; B chain; A chain; protease; fungal infection; malaria;  
XX leucocyte proliferation; cytomegalovirus; herpes; hepatitis;  
XX rhinovirus; laryngotracheitis; poliomyelitis; varicella zoster;  
XX cystic fibrosis; multiple sclerosis.  
XX Unidentified.  
XX Synthetic.

PN WO9849311-A2.  
 XX  
 PD  
 XX  
 PF  
 XX  
 PR  
 XX  
 30-APR-1998; 98WO-CA00394.  
 XX  
 PR  
 XX  
 29-OCT-1997; 97US-0063715.  
 PR  
 XX  
 30-APR-1997; 97US-0045148.  
 XX  
 PA  
 XX  
 (DNOV-) DE NOVO ENZYME CORP.  
 XX  
 PI  
 XX  
 Borgford T;  
 XX  
 WPI; 1999-009431/01.  
 XX  
 PT  
 XX  
 New nucleic acid encoding ricin-like toxin with an interchain linker  
 PT  
 PT  
 e.g. killing selectively cancer or infected cells  
 XX  
 PS  
 Claim 24; Fig 21; 352pp; English.  
 XX  
 CC  
 The present invention describes new purified and isolated nucleic acids  
 CC  
 (I) encoding: (i) the A and B chains of a ricin-like toxin (II); and  
 CC  
 (ii) a heterologous linker, joining the two chains and including a  
 CC  
 cleavage recognition site for a disease-specific protease (III). Also  
 CC  
 described are: (1) plasmids or baculovirus transfer vectors that contain  
 CC  
 (I); and (2) recombinant protein (IV) consisting of the A and B chains  
 CC  
 of (II) joined by the specified linker. (IV), produced by expression of  
 CC  
 (I) in host cells, are used to inhibit or kill diseased cells that  
 CC  
 produce (III), particularly for treating cancers (e.g. leucocyte  
 CC  
 proliferation; cancer of ovary, pancreas, breast or prostate; glioma) or  
 CC  
 infections caused by fungi, parasites (e.g. malaria) or viruses (e.g.  
 CC  
 cytomegalovirus (CMV), herpes, hepatitis, rhinovirus, laryngotracheitis,  
 CC  
 poliomyelitis or varicella zoster), also cystic fibrosis and multiple  
 CC  
 sclerosis. Alternatively, (I) is used to express (IV) in vivo. (IV) is  
 CC  
 toxic specifically for (III)-expressing cells and does not depend for  
 CC  
 specificity on a cell-binding component. When used to treat virus-  
 CC  
 infected cells, transcytosis and cytotoxicity of (IV) are increased by  
 CC  
 retrograde translocation from endoplasmic reticulum to cytoplasm (which  
 CC  
 some viruses exploit to avoid immune detection), so selectivity and  
 CC  
 safety are further improved. (IV) are not toxic until chain A is  
 CC  
 released and this occurs only in target cells. The present sequence  
 CC  
 represents a specifically claimed cancer protease-sensitive amino acid  
 CC  
 linker from the present invention.  
 XX  
 SQ  
 Sequence 12 AA;

Query Match 67.6%; Score 48; DB 20; Length 12;  
 Best Local Similarity 81.8%; Pred. No. 0.28;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
 |||||:|  
 Db 1 rpkpqgffglm 11

RESULT 96  
 AAG62769  
 ID AAG62769 standard; peptide; 12 AA.  
 XX  
 AC  
 XX  
 AAG62769;  
 XX  
 DT 17-SEP-2001 (first entry)  
 XX  
 DE Amino acid sequence of substance P precursor.  
 XX  
 KW Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200153336-A1.  
 XX

PD 26-JUL-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01529.  
 XX  
 PR 19-JAN-2000; 2000US-0489667.  
 XX  
 PA (ALLR ) ALLERGAN SALES INC.  
 XX  
 PI Donovan S;  
 XX  
 DR WPI; 2001-451900/48.  
 XX  
 PT Agent useful for treating pain comprises a clostridial neurotoxin (or  
 PT  
 component) attached to a targeting moiety  
 XX  
 PS Disclosure; Page 62; 77pp; English.  
 XX  
 CC The specification describes an agent, comprising a clostridial neurotoxin  
 CC  
 attached to a targeting moiety, where the targeting moiety is selected  
 CC  
 from transmission compounds, and compounds substantially similar to the  
 CC  
 transmission compounds. The agent may be used for treating pain, where  
 CC  
 the clostridial neurotoxin component is derived from botulinum toxin  
 CC  
 selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
 CC  
 The targeting moiety comprises a light chain and an amino end segment of  
 CC  
 a heavy chain and comprises Substance P as the targeting moiety. The pain  
 CC  
 alleviating effects persist for 2-6 months. The present sequence  
 CC  
 represents substance P precursor, and is used in the course of the  
 CC  
 invention.  
 XX  
 SQ Sequence 12 AA;

Query Match 67.6%; Score 48; DB 22; Length 12;  
 Best Local Similarity 81.8%; Pred. No. 0.28;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
 |||||:|  
 Db 1 rpkpqgffglm 11

RESULT 97  
 AAG62772  
 ID AAG62772 standard; peptide; 12 AA.  
 XX  
 AC  
 XX  
 AAG62772;  
 XX  
 DT 17-SEP-2001 (first entry)  
 XX  
 DE Amino acid sequence of carboxy-ester substance P precursor.  
 XX  
 KW Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FH Modified-site 12  
 FT /note= "methylated residue"  
 XX  
 PN WO200153336-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01529.  
 XX  
 PR 19-JAN-2000; 2000US-0489667.  
 XX  
 PA (ALLR ) ALLERGAN SALES INC.  
 XX  
 PI Donovan S;  
 XX  
 DR WPI; 2001-451900/48.  
 XX

PT Agent useful for treating pain comprises a clostridial neurotoxin (or  
PT component) attached to a targeting moiety -  
XX Disclosure; Page 64; 77pp; English.  
XX The specification describes an agent, comprising a clostridial neurotoxin  
CC attached to a targeting moiety, where the targeting moiety is selected  
CC from transmission compounds, and compounds substantially similar to the  
CC transmission compounds. The agent may be used for treating pain, where  
CC the clostridial neurotoxin component is derived from botulinum toxin  
CC selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
CC The targeting moiety comprises a light chain and an amine end segment of  
CC a heavy chain and comprises Substance P as the targeting moiety. The pain  
CC alleviating effects persist for 2-6 months. The present sequence  
CC represents a substance P precursor, and is used in the course of the  
CC invention.  
XX  
SQ Sequence 12 AA;  
  
Query Match 67.6%; Score 48; DB 22; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.28;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 RPKPQQWFWM 11  
| | | | | | | |  
Db 1 rpkpqffglm 11  
  
RESULT 98  
AAG62775  
ID AAG62775 standard; peptide; 12 AA.  
XX  
AC AAG62775;  
XX  
DT 17-SEP-2001 (first entry)  
XX  
DE Amino acid sequence of carboxy-ester substance P precursor.  
XX  
KW Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 12 /note= "ethylated residue"  
FT  
XX  
PN WO200153336-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01529.  
XX  
PR 19-JAN-2000; 2000US-0489667.  
XX  
PA (ALLR ) ALLERGAN SALES INC.  
XX  
PI Donovan S;  
XX  
DR WPI; 2001-451900/48.  
XX  
PT Agent useful for treating pain comprises a clostridial neurotoxin (or  
PT component) attached to a targeting moiety -  
XX Disclosure; Page 67; 77pp; English.  
XX  
CC The specification describes an agent, comprising a clostridial neurotoxin  
CC attached to a targeting moiety, where the targeting moiety is selected  
CC from transmission compounds, and compounds substantially similar to the  
CC transmission compounds. The agent may be used for treating pain, where  
CC the clostridial neurotoxin component is derived from botulinum toxin  
CC selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
CC The targeting moiety comprises a light chain and an amine end segment of

CC a heavy chain and comprises Substance P as the targeting moiety. The pain  
CC alleviating effects persist for 2-6 months. The present sequence  
CC represents a substance P precursor, and is used in the course of the  
CC invention.  
XX  
SQ Sequence 12 AA;  
  
Query Match 67.6%; Score 48; DB 22; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.28;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 RPKPQQWFWM 11  
| | | | | | | |  
Db 1 rpkpqffglm 11  
  
RESULT 99  
AAB84528  
ID AAB84528 standard; peptide; 12 AA.  
XX  
AC AAB84528;  
XX  
DT 05-SEP-2001 (first entry)  
XX  
DE Amino acid sequence of a modified substance P.  
XX  
KW Substance P; cell toxin; Pseudomonas exotoxin; cell ablation;  
KW NK-1 receptor; chronic pain; tumour; neurological dysfunction;  
KW basal ganglia; cholinergic interneuron; Parkinson's disease.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200131020-A1.  
XX  
PD 03-MAY-2001.  
XX  
PF 20-OCT-2000; 2000WO-US29064.  
XX  
PR 22-OCT-1999; 99US-0161159.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Fitzgerald DJ, Iadarola MJ;  
XX  
DR WPI; 2001-417560/44.  
XX  
PT Making cell toxin to treat chronic pain, by forming substance  
PT P-Pseudomonas exotoxin disulfide-linked conjugate, by reacting modified  
PT exotoxin and substance P having additional cysteine residue at its  
PT N-terminus -  
XX  
PS Example 1; Page 10; 54pp; English.  
XX  
CC The present sequence represents a modified substance P. The peptide is  
CC used to produce a cell toxin. The cell toxin comprises a substance  
CC P-Pseudomonas exotoxin disulfide-linked conjugate. The cell toxin is  
CC useful for ablating NK-1 receptor expressing cells, such as dorsal horn  
CC cell, a stratum cell or a brain parenchyma cell, for treating chronic  
CC pain in epineurium cells, perineurium cells, nerve ganglia, nerve  
CC sheaths, nerve linings, meninges, pia mater cells, arachnoid membrane  
CC cells, duramembrane cells, cells lining a joint or brain or spinal cord  
CC parenchymal cells, without significantly affecting basal nociceptive  
CC responses. The cell toxin is thus useful for treating chronic pain or  
CC tumours that binds substance P. It is also useful for neurological  
CC dysfunctions of the basal ganglia by targeting cholinergic interneurons  
CC that express substance P e.g. Parkinson's disease.  
XX  
SQ Sequence 12 AA;  
  
Query Match 67.6%; Score 48; DB 22; Length 12;

```
Best Local Similarity 81.8%; Pred. No. 0.28;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11
Db 2 rpkpqffglm 12

RESULT 100
AAB98867
ID AAB98867 standard; Peptide; 12 AA.
XX AC AAB98867;
XX XX 14-AUG-2001 (first entry)
XX XX
DE Chimeric analgesic peptide #23.
XX XX
KW Opioid receptor binding; nociceptive receptor binding; analgesic;
KW pain; chimeric peptide.
XX XX
OS Synthetic.
XX XX
FH Key Location/Qualifiers
FT Modified-site 12
FT /label= OTHER
FT /note= "modified by Ome"
XX XX
PN WO200130371-A2.
XX XX
PD 03-MAY-2001.
XX XX
PF 27-OCT-2000; 2000WO-US29789.
XX XX
PR 28-OCT-1999; 99US-0428692.
XX XX
PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.
XX XX
PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;
XX XX
DR WPI; 2001-397593/42.
XX XX
PT New chimeric peptides used for treating pain comprise opioid receptor
FT binding group and nociceptive receptor binding group
XX XX
PS Claim 10; Page 15; 34pp; English.
XX XX
CC The present invention describes a number of chimeric peptides comprising
CC an opioid receptor binding moiety and a nociceptive receptor binding
CC moiety. These can be used as analgesics for the treatment of pain. Unlike
CC opioid receptor based peptides alone, tolerance does not result from
CC their long-term use. The present sequence is one of the peptides of the
CC invention.
XX XX
SQ Sequence 12 AA;

Query Match 67.6%; Score 48; DB 22; Length 12;
Best Local Similarity 81.8%; Pred. No. 0.28;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11
Db 2 rpkpqffglm 12

RESULT 102
AAB98873
ID AAB98873 standard; Peptide; 12 AA.
XX AC AAB98873;
XX XX
DT 14-AUG-2001 (first entry)
XX XX
DE Chimeric analgesic peptide #29.
XX XX
KW Opioid receptor binding; nociceptive receptor binding; analgesic;
KW pain; chimeric peptide.
XX XX
OS Synthetic.
XX XX
FH Key Location/Qualifiers
FT Modified-site 12
FT /label= OTHER
FT /note= "modified by Oeth"
XX XX
PN WO200130371-A2.
XX XX

Best Local Similarity 81.8%; Pred. No. 0.28;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11
Db 2 rpkpqffglm 12

RESULT 101
AAB98870
ID AAB98870 standard; Peptide; 12 AA.
XX AC AAB98870;
XX XX
DT 14-AUG-2001 (first entry)
```



PD 03-MAY-2001. PD  
 XX  
 PF 27-OCT-2000; 2000WO-US29789. PF  
 XX  
 PR 28-OCT-1999; 99US-0428692. PR  
 XX  
 PA (NEWF-) NEW ENGLAND MEDICAL CENT HOSPITALS INC. PA  
 XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A; XX  
 XX WPI; 2001-397593/42. XX  
 DR  
 XX New chimeric peptides used for treating pain comprise opioid receptor  
 PT binding group and nociceptive receptor binding group  
 XX  
 PS Claim 10; Page 15; 34pp; English. PS  
 XX The present invention describes a number of chimeric peptides comprising  
 CC an opioid receptor binding moiety and a nociceptive receptor binding  
 CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
 CC opioid receptor based peptides alone, tolerance does not result from  
 CC their long-term use. The present sequence is one of the peptides of the  
 CC invention.  
 XX  
 SQ Sequence 12 AA;  
 Query Match 67.6%; Score 48; DB 22; Length 12;  
 Best Local Similarity 81.8%; Pred. No. 0.28;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RPKPQOWFWLM 11  
 Db |||||:| ||  
 1 rpkpqffgilm 11  
 RESULT 103  
 AAY03158  
 ID AAY03158 standard; peptide; 13 AA.  
 XX  
 AC AAY03158;  
 XX  
 DT 10-JUN-1999 (first entry)  
 XX  
 DE Substance P-Glycine-Lysine.  
 XX  
 KW Opioid peptide; opioid analgesia; enhancer; opioid anaesthesia;  
 XX substance P.  
 XX  
 OS Synthetic.  
 XX  
 PN US5891842-A.  
 XX  
 PD 06-APR-1999.  
 XX  
 PF 12-APR-1996; 96US-0631434.  
 XX  
 PR 09-APR-1993; 93US-0044954.  
 XX  
 PR 12-APR-1996; 96US-0631434.  
 XX  
 PA (TUFT ) TUFTS COLLEGE.  
 XX  
 PI Kream RM;  
 XX  
 DR WPI; 1999-253906/21.  
 XX  
 XX Synergistic method for enhancing opioid analgesia and anaesthesia  
 PT within a human  
 XX  
 PS Disclosure; Column 14; 20pp; English.  
 XX  
 CC This sequence represents substance P used in the method of the  
 CC invention. The method is for enhancing opioid analgesia within a human

CC subject for a duration of 15 minutes comprises concurrent administration  
 CC of substance P, or one of its precursors. The method is used to elicit  
 CC opioid analgesia and anaesthesia, either prior to or after the occurrence  
 CC of a nociceptive event. The components have a synergistic effect. The  
 CC method allows use of low doses of opioid that produce little or no  
 CC physiological effect reducing conventional risks of toxicity and  
 CC addiction, and allows the use of low doses of substance P and its related  
 CC analogs that limit their in vivo physiological consequences. The  
 CC analgesia is naloxone reversible allowing diminishment or complete  
 CC elimination of opioid analgesia if desired and on demand. The treatment  
 CC provides a durable analgesic effect, but only minimally disturbs and  
 CC interrupts the normal metabolic processes of the body.  
 XX  
 SQ Sequence 13 AA;  
 Query Match 67.6%; Score 48; DB 20; Length 13;  
 Best Local Similarity 81.8%; Pred. No. 0.3;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RPKPQOWFWLM 11  
 Db |||||:| ||  
 1 rpkpqffgilm 11  
 RESULT 104  
 AAG62770  
 ID AAG62770 standard; peptide; 13 AA.  
 XX  
 AC AAG62770;  
 XX  
 DT 17-SEP-2001 (first entry)  
 XX  
 DE Amino acid sequence of substance P precursor.  
 XX  
 KW Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200153336-A1.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01529.  
 XX  
 PR 19-JAN-2000; 2000US-0489667.  
 XX  
 PA (ALLR ) ALLERGAN SALES INC.  
 XX  
 PI Donovan S;  
 XX  
 DR WPI; 2001-451900/48.  
 XX  
 PT Agent useful for treating pain comprises a clostridial neurotoxin (or  
 PT component) attached to a targeting moiety  
 XX  
 PS Disclosure; Page 62; 77pp; English.  
 XX  
 CC The specification describes an agent, comprising a clostridial neurotoxin  
 CC attached to a targeting moiety, where the targeting moiety is selected  
 CC from transmission compounds, and compounds substantially similar to the  
 CC transmission compounds. The agent may be used for treating pain, where  
 CC the clostridial neurotoxin component is derived from botulinum toxin  
 CC selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
 CC The targeting moiety comprises a light chain and an amine end segment of  
 CC a heavy chain and comprises Substance P as the targeting moiety. The pain  
 CC alleviating effects persist for 2-6 months. The present sequence  
 CC represents substance P precursor, and is used in the course of the  
 CC invention.  
 XX  
 SQ Sequence 13 AA;

Query Match 67.6%; Score 48; DB 22; Length 13;  
Best Local Similarity 81.8%; Pred. No. 0.3;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 1 rpkpqgfglm 11

## RESULT 105

AAG62773  
ID AAG62773 standard; peptide; 13 AA.

XX AC AAG62773;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of carboxy-ester substance P precursor.

KW Clostridial neurotoxin; pain; botulinum toxin; Substance P.

XX OS Synthetic.

XX FH Key Location/Qualifiers  
FT Modified-site 13  
FT /note= "methylated residue"

XX PN WO200153336-A1.

XX PD 26-JUL-2001.

XX PF 17-JAN-2001; 2001WO-US01529.

XX PR 19-JAN-2000; 2000US-0489667.

XX PA (ALLR ) ALLERGAN SALES INC.

XX PI Donovan S;

XX DR WPI; 2001-451900/48.

XX PT Agent useful for treating pain comprises a clostridial neurotoxin (or  
PT component) attached to a targeting moiety -

XX PS Disclosure; Page 65; 77pp; English.

XX CC The specification describes an agent, comprising a clostridial neurotoxin  
CC attached to a targeting moiety, where the targeting moiety is selected  
CC from transmission compounds, and compounds substantially similar to the  
CC transmission compounds. The agent may be used for treating pain, where  
CC the clostridial neurotoxin component is derived from botulinum toxin  
CC selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
CC The targeting moiety comprises a light chain and an amine end segment of  
CC a heavy chain and comprises Substance P as the targeting moiety. The pain  
CC alleviating effects persist for 2-6 months. The present sequence  
CC represents a substance P precursor, and is used in the course of the  
CC invention.

XX SQ Sequence 13 AA;

Query Match 67.6%; Score 48; DB 22; Length 13;  
Best Local Similarity 81.8%; Pred. No. 0.3;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 1 rpkpqgfglm 11

## RESULT 106

AAG62776  
ID AAG62776 standard; peptide; 13 AA.

FH Key Location/Qualifiers  
FT Modified-site 13  
FT /label= OTHER  
FT /note= "C-terminal amide"  
XX  
XX WO200130371-A2.  
XX  
XX 03-MAY-2001.  
XX  
XX 27-OCT-2000; 2000WO-US29789.  
XX  
XX 28-OCT-1999; 99US-0428692.  
XX  
XX (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX  
XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX WPI; 2001-397593/42.  
XX  
XX New chimeric peptides used for treating pain comprise opioid receptor binding group and nociceptive receptor binding group  
XX  
XX Claim 10; Page 15; 34pp; English.  
XX  
XX The present invention describes a number of chimeric peptides comprising an opioid receptor binding moiety and a nociceptive receptor binding moiety. These can be used as analgesics for the treatment of pain. Unlike opioid receptor based peptides alone, tolerance does not result from their long-term use. The present sequence is one of the peptides of the invention.  
XX  
XX SQ Sequence 13 AA;  
XX  
XX Query Match 67.6%; Score 48; DB 22; Length 13;  
XX Best Local Similarity 81.8%; Pred. No. 0.3;  
XX Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX QY 1 RPKPQQWFWM 11  
XX |||||: ||  
XX Db 1 rpkpqffglm 11  
XX  
XX RESULT 108  
XX AAB98871  
XX ID AAB98871 standard; Peptide; 13 AA.  
XX AC AAB98871;  
XX  
XX DT 14-AUG-2001 (first entry)  
XX  
XX DE Chimeric analgesic peptide #27.  
XX  
XX KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
XX KW pain; chimeric peptide.  
XX OS Synthetic.  
XX  
XX FH Key Location/Qualifiers  
XX FT Modified-site 13  
XX FT /label= OTHER  
XX FT /note= "modified by Ome"  
XX  
XX PN WO200130371-A2.  
XX  
XX PD 03-MAY-2001.  
XX  
XX PF 27-OCT-2000; 2000WO-US29789.  
XX  
XX PR 28-OCT-1999; 99US-0428692.  
XX  
XX PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX  
XX PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX DR WPI; 2001-397593/42.  
XX  
XX PT New chimeric peptides used for treating pain comprise opioid receptor binding group and nociceptive receptor binding group  
XX  
XX PS Claim 10; Page 15; 34pp; English.  
XX  
XX CC The present invention describes a number of chimeric peptides comprising an opioid receptor binding moiety and a nociceptive receptor binding moiety. These can be used as analgesics for the treatment of pain. Unlike opioid receptor based peptides alone, tolerance does not result from their long-term use. The present sequence is one of the peptides of the invention.  
XX  
XX SQ Sequence 13 AA;  
XX  
XX Query Match 67.6%; Score 48; DB 22; Length 13;  
XX Best Local Similarity 81.8%; Pred. No. 0.3;  
XX Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX QY 1 RPKPQQWFWM 11  
XX |||||: ||  
XX Db 1 rpkpqffglm 11  
XX  
XX RESULT 108  
XX AAB98871  
XX ID AAB98871 standard; Peptide; 13 AA.  
XX AC AAB98871;  
XX  
XX DT 14-AUG-2001 (first entry)  
XX  
XX DE Chimeric analgesic peptide #27.  
XX  
XX KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
XX KW pain; chimeric peptide.  
XX OS Synthetic.  
XX  
XX FH Key Location/Qualifiers  
XX FT Modified-site 13  
XX FT /label= OTHER  
XX FT /note= "modified by Ome"  
XX  
XX PN WO200130371-A2.  
XX  
XX PD 03-MAY-2001.  
XX  
XX PF 27-OCT-2000; 2000WO-US29789.  
XX  
XX PR 28-OCT-1999; 99US-0428692.  
XX  
XX PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX  
XX WPI; 2001-397593/42.  
XX  
XX New chimeric peptides used for treating pain comprise opioid receptor binding group and nociceptive receptor binding group  
XX  
XX Claim 10; Page 15; 34pp; English.  
XX  
XX The present invention describes a number of chimeric peptides comprising an opioid receptor binding moiety and a nociceptive receptor binding moiety. These can be used as analgesics for the treatment of pain. Unlike opioid receptor based peptides alone, tolerance does not result from their long-term use. The present sequence is one of the peptides of the invention.  
XX  
XX SQ Sequence 13 AA;  
XX  
XX Query Match 67.6%; Score 48; DB 22; Length 13;  
XX Best Local Similarity 81.8%; Pred. No. 0.3;  
XX Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX QY 1 RPKPQQWFWM 11  
XX |||||: ||  
XX Db 1 rpkpqffglm 11  
XX  
XX RESULT 109  
XX AAB98874  
XX ID AAB98874 standard; Peptide; 13 AA.  
XX AC AAB98874;  
XX  
XX DT 14-AUG-2001 (first entry)  
XX  
XX DE Chimeric analgesic peptide #30.  
XX  
XX KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
XX KW pain; chimeric peptide.  
XX OS Synthetic.  
XX  
XX FH Key Location/Qualifiers  
XX FT Modified-site 13  
XX FT /label= OTHER  
XX FT /note= "modified by Oeth"  
XX  
XX PN WO200130371-A2.  
XX  
XX PD 03-MAY-2001.  
XX  
XX PF 27-OCT-2000; 2000WO-US29789.  
XX  
XX PR 28-OCT-1999; 99US-0428692.  
XX  
XX PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX  
XX PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX DR WPI; 2001-397593/42.  
XX  
XX PT New chimeric peptides used for treating pain comprise opioid receptor binding group and nociceptive receptor binding group  
XX  
XX PS Claim 10; Page 15; 34pp; English.  
XX  
XX CC The present invention describes a number of chimeric peptides comprising an opioid receptor binding moiety and a nociceptive receptor binding moiety. These can be used as analgesics for the treatment of pain. Unlike opioid receptor based peptides alone, tolerance does not result from their long-term use. The present sequence is one of the peptides of the invention.  
XX  
XX CC

XX  
SQ Sequence 13 AA;

Query Match 67.6%; Score 48; DB 22; Length 13;  
Best Local Similarity 81.8%; Pred. No. 0.32;  
Matches 9; Conservative 1; Mismatches 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|  
Db 1 rpkpqgffglm 11

RESULT 110

AAY03159

ID AAY03159 standard; peptide; 14 AA.

XX

AC AAY03159;

XX 10-JUN-1999 (first entry)

XX Substance P-Glycine-Lysine-Arginine.

DE  
XX  
XX Opioid peptide; opioid analgesia; enhancer; opioid anaesthesia;  
KW substance P.  
XX  
XX Synthetic.  
OS  
XX US5891842-A.

PN

XX 06-APR-1999.

XX 12-APR-1996; 96US-0631434.

XX 09-APR-1993; 93US-0044954.

PR 12-APR-1996; 96US-0631434.

XX (TUFT ) TUFTS COLLEGE.

XX Kream RM;

XX WPI; 1999-253906/21.

DR Synergistic method for enhancing opioid analgesia and anaesthesia  
XX within a human  
XX Disclosure; Column 14; 20pp; English.

PS This sequence represents substance P used in the method of the

XX invention. The method is for enhancing opioid analgesia within a human  
CC subject for a duration of 15 minutes comprises concurrent administration  
CC of substance P, or one of its precursors. The method is used to elicit  
CC opioid analgesia and anaesthesia, either prior to or after the occurrence  
CC of a nociceptive event. The components have a synergistic effect. The  
CC method allows use of low doses of opioid that produce little or no  
CC physiological effect reducing conventional risks of toxicity and  
CC addiction, and allows the use of low doses of substance P and its related  
CC analogs that limit their in vivo physiological consequences. The  
CC analgesia is naloxone reversible allowing diminishment or complete  
CC elimination of opioid analgesia if desired and on demand. The treatment  
CC provides a durable analgesic effect, but only minimally disturbs and  
CC interrupts the normal metabolic processes of the body.

XX Sequence 14 AA;

Query Match 67.6%; Score 48; DB 20; Length 14;  
Best Local Similarity 81.8%; Pred. No. 0.32;  
Matches 9; Conservative 1; Mismatches 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|  
Db 1 rpkpqgffglm 11

RESULT 111

AAG62771

ID AAG62771 standard; peptide; 14 AA.

XX

AC AAG62771;

XX 17-SEP-2001 (first entry)

XX Amino acid sequence of substance P precursor.

DE Clostridial neurotoxin; pain; botulinum toxin; Substance P.

XX Unidentified.

XX WO200153336-A1.

XX 26-JUL-2001.

XX 17-JAN-2001; 2001WO-US01529.

XX 19-JAN-2000; 2000US-0489667.

XX (ALLR ) ALLERGAN SALES INC.

XX Donovan S;

XX WPI; 2001-451900/48.

XX Agent useful for treating pain comprises a clostridial neurotoxin (or

XX component) attached to a targeting moiety -

XX Disclosure; Page 63; 77pp; English.

XX The specification describes an agent, comprising a clostridial neurotoxin  
CC attached to a targeting moiety, where the targeting moiety is selected  
CC from transmission compounds, and compounds substantially similar to the  
CC transmission compounds. The agent may be used for treating pain, where  
CC the clostridial neurotoxin component is derived from botulinum toxin  
CC selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
CC The targeting moiety comprises a light chain and an amine end segment of  
CC a heavy chain and comprises substance P as the targeting moiety. The pain  
CC alleviating effects persist for 2-6 months. The present sequence  
CC represents substance P precursor, and is used in the course of the  
CC invention.

XX Sequence 14 AA;

Query Match 67.6%; Score 48; DB 22; Length 14;  
Best Local Similarity 81.8%; Pred. No. 0.32;  
Matches 9; Conservative 1; Mismatches 0; Gaps 0;

QY 1 RPKPQQWFWM 11

|||||:|

Db 1 rpkpqgffglm 11

RESULT 112

AAG62774

ID AAG62774 standard; peptide; 14 AA.

XX

AC AAG62774;

XX 17-SEP-2001 (first entry)

XX Amino acid sequence of carboxy-ester substance P precursor.

DE Clostridial neurotoxin; pain; botulinum toxin; Substance P.

XX Synthetic.

OS

XX

Key Location/Qualifiers  
Modified-site 14  
/note= "methylated residue"

WO200153336-A1.

26-JUL-2001.

17-JAN-2001; 2001WO-US01529.

19-JAN-2000; 2000US-0489667.

(ALLR ) ALLERGAN SALES INC.

Donovan S;

WPI; 2001-451900/48.

Agent useful for treating pain comprises a clostridial neurotoxin (or component) attached to a targeting moiety -

Disclosure; Page 66; 77pp; English.

The specification describes an agent, comprising a clostridial neurotoxin attached to a targeting moiety, where the targeting moiety is selected from transmission compounds, and compounds substantially similar to the transmission compounds. The agent may be used for treating pain, where the clostridial neurotoxin component is derived from botulinum toxin selected from botulinum types A, B, C, D, E, F, G and mixtures of these. The targeting moiety comprises a light chain and an amine end segment of a heavy chain and comprises Substance P as the targeting moiety. The pain alleviating effects persist for 2-6 months. The present sequence represents a substance P precursor, and is used in the course of the invention.

Sequence 14 AA;

Query Match 67.6%; Score 48; DB 22; Length 14;  
Best Local Similarity 81.8%; Pred. No. 0.32;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 rpkpqffglm 11

RESULT 113

AAG62777  
ID AAG62777 standard; peptide; 14 AA.

AC AAG62777;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of carboxy-ester substance P precursor.

KW Clostridial neurotoxin; pain; botulinum toxin; Substance P.

XX Synthetic.

Key Location/Qualifiers  
Modified-site 14  
/note= "ethylated residue"

WO200153336-A1.

26-JUL-2001.

17-JAN-2001; 2001WO-US01529.

19-JAN-2000; 2000US-0489667.

PA (ALLR ) ALLERGAN SALES INC.

PI Donovan S;

XX WPI; 2001-451900/48.

XX Agent useful for treating pain comprises a clostridial neurotoxin (or component) attached to a targeting moiety -

XX Disclosure; Page 69; 77pp; English.

XX The specification describes an agent, comprising a clostridial neurotoxin attached to a targeting moiety, where the targeting moiety is selected from transmission compounds, and compounds substantially similar to the transmission compounds. The agent may be used for treating pain, where the clostridial neurotoxin component is derived from botulinum toxin selected from botulinum types A, B, C, D, E, F, G and mixtures of these. The targeting moiety comprises a light chain and an amine end segment of a heavy chain and comprises Substance P as the targeting moiety. The pain alleviating effects persist for 2-6 months. The present sequence represents a substance P precursor, and is used in the course of the invention.

XX Sequence 14 AA;

Query Match 67.6%; Score 48; DB 22; Length 14;  
Best Local Similarity 81.8%; Pred. No. 0.32;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 rpkpqffglm 11

RESULT 114

AAB98869  
ID AAB98869 standard; Peptide; 14 AA.

AC AAB98869;

DT 14-AUG-2001 (first entry)

DE Chimeric analgesic peptide #25.

XX Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.

XX Synthetic.

Key Location/Qualifiers  
Modified-site 14  
/label= OTHER  
/note= "C-terminal amide"

XX WO200130371-A2.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-US29789.

XX 28-OCT-1999; 99US-0428692.

XX (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;

XX WPI; 2001-397593/42.

XX New chimeric peptides used for treating pain comprise opioid receptor binding group and nociceptive receptor binding group -

XX Claim 10; Page 15; 34pp; English.

XX The present invention describes a number of chimeric peptides comprising  
CC an opioid receptor binding moiety and a nociceptive receptor binding  
CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
CC opioid receptor based peptides alone, tolerance does not result from  
CC their long-term use. The present sequence is one of the peptides of the  
CC invention.  
XX  
SQ Sequence 14 AA;

Query Match 67.6%; Score 48; DB 22; Length 14;  
Best Local Similarity 81.8%; Pred. No. 0.32;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
Db 1 rpkpqffgim 11  
|||||:| |

RESULT 115  
AAB98872  
ID AAB98872 standard; Peptide; 14 AA.  
XX  
AC AAB98872;  
XX  
DT 14-AUG-2001 (first entry)  
XX  
DE Chimeric analgesic peptide #28.  
XX  
KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 14  
FT /label= OTHER  
FT /note= "modified by OMe"  
XX  
PN WO200130371-A2.  
XX  
PD 03-MAY-2001.  
XX  
PF 27-OCT-2000; 2000WO-US29789.  
XX  
PR 28-OCT-1999; 99US-0428692.  
XX  
PA (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX  
PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX  
DR WPI; 2001-397593/42.  
XX  
PT New chimeric peptides used for treating pain comprise opioid receptor  
binding group and nociceptive receptor binding group  
XX  
PS Claim 10; Page 15; 34pp; English.  
XX  
CC The present invention describes a number of chimeric peptides comprising  
CC an opioid receptor binding moiety and a nociceptive receptor binding  
CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
CC opioid receptor based peptides alone, tolerance does not result from  
CC their long-term use. The present sequence is one of the peptides of the  
CC invention.  
XX  
SQ Sequence 14 AA;

Query Match 67.6%; Score 48; DB 22; Length 14;  
Best Local Similarity 81.8%; Pred. No. 0.32;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
Db 1 rpkpqffgim 11  
|||||:| |

RESULT 116  
AAB98875  
ID AAB98875 standard; Peptide; 14 AA.  
XX  
AC AAB98875;  
XX  
DT 14-AUG-2001 (first entry)  
XX  
DE Chimeric analgesic peptide #31.  
XX  
KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 14  
FT /label= OTHER  
FT /note= "modified by Oeth"  
XX  
PN WO200130371-A2.  
XX  
PD 03-MAY-2001.  
XX  
PF 27-OCT-2000; 2000WO-US29789.  
XX  
PR 28-OCT-1999; 99US-0428692.  
XX  
PA (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX  
PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX  
DR WPI; 2001-397593/42.  
XX  
PT New chimeric peptides used for treating pain comprise opioid receptor  
binding group and nociceptive receptor binding group  
XX  
PS Claim 10; Page 15; 34pp; English.  
XX  
CC The present invention describes a number of chimeric peptides comprising  
CC an opioid receptor binding moiety and a nociceptive receptor binding  
CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
CC opioid receptor based peptides alone, tolerance does not result from  
CC their long-term use. The present sequence is one of the peptides of the  
CC invention.  
XX  
SQ Sequence 14 AA;

Query Match 67.6%; Score 48; DB 22; Length 14;  
Best Local Similarity 81.8%; Pred. No. 0.32;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
Db 1 rpkpqffgim 11  
|||||:| |

RESULT 117  
AAB91440  
ID AAB91440 standard; Peptide; 14 AA.  
XX  
AC AAB91440;  
XX  
DT 22-JUN-2001 (first entry)  
XX  
DE Tachykinins peptide SEQ ID NO:616.  
XX  
SQ

KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX Homo sapiens.  
OS Synthetic.  
XX WO200069900-A2.  
XX 23-NOV-2000.  
XX 17-MAY-2000; 2000WO-US13576.  
XX 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX (CONJ-) CONJUCHEM INC.  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
PI WPI; 2001-112059/12.  
XX  
XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
XX Disclosure; Page 400; 733pp; English.  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides with physiological processes.  
XX exemplification of the present invention.  
XX  
XX Sequence 14 AA;  
Query Match 67.6%; Score 48; DB 22; Length 14;  
Best Local Similarity 81.8%; Pred. No. 0.32;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
DB 1 rpkpqffgilm 11  
RESULT 118  
AAB06258  
ID AAB06258 standard; peptide; 20 AA.  
XX AAB06258;  
XX 16-OCT-2000 (first entry)  
XX Substance P analogue #2.  
XX  
XX Substance P; SP; neurokinin-1 receptor; NK-1R; nociception; NTE-SAP;  
KW saporin; SAP; analgesic; anti-inflammatory; neuroprotective;  
KW anti-asthmatic; anti-allergic; dermatological; anti-ulcer;  
KW tranquiliser; immunosuppressive; anti-migraine; cytostatic;

KW substance P antagonist; cytotoxic; ribosome inactivator;  
KW prostaglandin antagonist; cancer; respiratory disease; asthma;  
KW allergic rhinitis; ophthalmic disease; conjunctivitis;  
KW allergic dermatitis; psoriasis; ulcerative colitis; Crohn's disease;  
KW gastrointestinal disorder; anxiety; psychosis; rheumatoid arthritis;  
KW carcinoma; lupus erythematosus conjunctivitis.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 20  
FT /note= "C-terminal amide"  
XX  
XX US6063758-A.  
XX 16-MAY-2000.  
XX  
XX 09-JUL-1997; 97US-0890157.  
XX  
XX 09-JUL-1997; 97US-0890157.  
XX (ADTA-) ADVANCED TARGETING SYSTEMS INC.  
XX  
XX Lappi DA, Wiley RG;  
XX WPI; 2000-430049/37.  
XX  
XX New conjugates comprising substance P or its analog, and a  
PT ribosome-inactivating protein (for example saporin), for alleviating  
PT pain and treating disorders associated with neurokinin-1 receptor -  
XX Claim 1; Column 2; 21pp; English.  
XX  
XX The present sequence is an analogue of substance P (SP). SP, which binds  
CC to the neurokinin-1 receptor (NK-1R), is best known for its role in  
CC nociception. It is secreted by small unmyelinated C-fibres of the  
CC peripheral nervous system that are thought to be primary nociceptive  
CC neurons. The present sequence may be conjugated to Saporin (SAP), a  
CC ribosome-inactivating protein, to produce NTE-SAP. The conjugate may be  
CC used to control chronic pain by specifically targeting cells having NK1  
CC receptors, and inhibiting proliferation of or causing death of these  
CC cells. It may also be used to treat NK-1R-associated disorders  
CC including respiratory conditions (e.g. asthma, allergic rhinitis),  
CC ophthalmic conditions (e.g. conjunctivitis), cutaneous conditions (e.g.  
CC allergic dermatitis, psoriasis), intestinal conditions (e.g. ulcerative  
CC colitis, Crohn's disease), gastrointestinal disorders, central nervous  
CC system disorders (e.g. anxiety, psychosis), inflammatory diseases (e.g.  
CC rheumatoid arthritis), proliferative conditions (e.g. carcinoma),  
CC disorders related to immune enhancement or suppression (e.g. lupus  
CC erythematosus conjunctivitis), and especially migraine.  
XX  
XX Sequence 20 AA;  
Query Match 67.6%; Score 48; DB 21; Length 20;  
Best Local Similarity 81.8%; Pred. No. 0.45;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
DB 10 rpkpqffgilm 20  
RESULT 119  
AAP70431  
ID AAP70431 standard; protein; 129 AA.  
XX AAP70431;  
XX  
XX 17-JAN-1991 (first entry)  
DT Human beta-preprotachykinin.  
XX

KW Preprotachykinin; substance P; neurokinin A; tachykinin;

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Region 20..56

XX FT /label=claimed polypeptide

XX FT Region 1..126

XX FT /label=claimed polypeptide

XX FT Region 111..126

XX FT /label=claimed polypeptide

XX PN WO8707643-A.

XX PD 17-DEC-1987.

XX PF 03-JUN-1987; 87WO-GB00382.

XX PR 03-JUN-1986; 86GB-0013431.

XX PA (RESE ) RESEARCH CORPORATION LTD.

XX PI Harnar AJ, Pascall J, McKeown A;

XX DR WPI; 1987-362730/51.

XX DR N-PSDB; AAN70688.

XX PT New DNA sequence coding for the new polypeptide preprotachykinin -

XX PT a precursor for substance P, etc., useful as neurotransmitters,

XX PT diagnostic reagents, etc.

XX PS Claim 1; page 15; 25pp; English.

XX CC Beta-preprotachykinin includes sequences identical to tachykinins, eg

XX CC substance P, neurokinin A, or other biologically active peptides, eg

XX CC neuropeptide K. These peptides are, eg neurotransmitters, hormones,

XX CC analgesics and anti-inflammatories. The polypeptides can be used

XX CC as reagents in RIA, eg to monitor or diagnose carcinoid syndrome.

XX SQ Sequence 129 AA;

XX SQ Sequence 129 AA;

Query Match

Best Local Similarity 67.6%; Score 48; DB 8; Length 129;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOWFWLM 11

DB 58 rpqpgffglm 68

RESULT 120

ID AAG99353 standard; Protein; 129 AA.

XX AC AAG99353;

XX DT 25-SEP-2001 (first entry)

XX DE Human atypical tachykinin protein fragment SEQ ID NO: 63.

XX KW Atypical tachykinin; ATT; human; hypertension.

XX OS Homo sapiens.

XX PN WO200146415-A1.

XX PD 28-JUN-2001.

XX PF 21-DEC-2000; 2000WO-JP09083.

XX PR 21-DEC-1999; 99JP-0362638.

XX PR 10-MAR-2000; 2000JP-0066714.

XX PA (TAKE ) TAKEDA CHEM IND LTD.

XX PI Itoh Y, Nishi K, Kitada C, Inatomi N;

XX DR WPI; 2001-441676/47.

XX PT Atypical tachykinin peptides of human origin and DNA encoding them for

XX PT screening potential agents for treatment of hypertension -

XX PS Example 14; Page 143; 153pp; Japanese.

XX CC The present invention relates to atypical tachykinin proteins of human

XX CC origin and their esters, amides, salts and partial peptides. These can be

XX CC used in the treatment, prevention and diagnosis of hypertension. The

XX CC present sequence is a protein fragment described in the exemplification

XX CC of the invention.

XX SQ Sequence 129 AA;

Query Match 67.6%; Score 48; DB 22; Length 129;

Best Local Similarity 81.8%; Pred. No. 2.7;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOWFWLM 11

DB 58 rpqpgffglm 68

RESULT 121

ID AAW16339 standard; Protein; 401 AA.

XX AC AAW16339;

XX DT 05-SEP-1997 (first entry)

XX DE DAB389-SP-Gly fusion toxin.

XX KW DAB389-SP-Gly; amidated polypeptide binding ligand; drug delivery;

XX KW diphtheria toxin; substance P; cancer; therapy.

XX OS Synthetic.

XX PN WO9713410-A1.

XX PD 17-APR-1997.

XX PF 11-OCT-1996; 96WO-US16237.

XX PR 13-OCT-1995; 95US-0005431.

XX PA (BOST-) BOSTON MEDICAL CENT CORP.

XX PI Fisher CE, Leeman SE, Murphy JR, Vanderspek JC;

XX DR WPI; 1997-235583/21.

XX DR N-PSDB; AAT63359.

XX PT Hybrid molecule for targeting compound, especially a toxin, into

XX PT cells - includes polypeptide able to transport the compound across

XX PT cytoplasmic membranes and amidated ligand, useful for treatment of

XX PT cancer

XX PS Example 1; Page 22-23; 51pp; English.

XX CC DAB389-SP-Gly (AAW16339) is a hybrid toxin comprising DAB389 (i.e.

XX CC amino acids 1-386 plus His-484 and Ala-485 of mature diphtheria

XX CC toxin) fused to C-terminal glycine-extended substance P. It was

XX CC expressed in E. coli HMS174(DE3) transformants using a vector

XX CC that carried DAB389-SP-Gly DNA (see also AAT63359). The fusion

XX CC protein was then amidated using peptidylglycino-alpha-amidating



CC monoxigenase. The amidated fusion protein used to target DAB389  
 CC toxin to specific cells contg. substance P receptors, esp. cancer  
 CC cells. For human IM9 (chronic myelogenous leukaemia) cells contg.  
 CC approx. 4000 substance P receptors per cell, the IC50 for amidated  
 CC DAB389-SP-Gly was 18 pM.

XX Sequence 401 AA;

Query Match 67.6%; Score 48; DB 18; Length 401;  
 Best Local Similarity 81.8%; Pred. No. 8.1;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOWFWLM 11  
 |||||:|  
 Db 390 rpkpqffgilm 400

RESULT 122

AAW26510  
 ID AAW26510 standard; Protein; 487 AA.

XX AC AAW26510;

XX DT 06-JAN-1998 (first entry)

XX DE Amyloid precursor protein substrate APP-REP 751.

XX KW Amyloid precursor protein; APP; beta-amyloid protein; BAP;  
 KW substrate; mutein; secretase; Alzheimer's disease; human;  
 KW APP-REP 751; pCLL621.

XX OS Chimeric Homo sapiens.  
 OS Chimeric synthetic.

XX FH Key Location/Qualifiers

FT Peptide 362..372  
 FT /label= SP  
 FT /note= "substance P reporter epitope"

FT Domain 389..430

FT /label= BAP

FT /note= "beta-amyloid protein"

FT Cleavage-site 404..405

FT /note= "secretase cleavage site"

FT Domain 417..440

FT /label= Transmembrane

XX US5656477-A.

PN 12-AUG-1997.

PD 01-MAY-1992; 92US-0877675.

PF 20-SEP-1993; 93US-0123659.

PR 01-MAY-1992; 92US-0877675.

XX (AMCY ) AMERICAN CYANAMID CO.

XX Jacobsen JS, Vitek MP;

XX WPI; 1997-414594/38.

XX P-PSDB; AAT87083.

XX Nucleic acid encoding amyloid precursor mutein(s) - comprising  
 PT reporter gene and coding sequence, for identifying compounds which  
 PT modify the activity of proteolytic enzymes which cleave APP  
 XX Disclosure; Fig 8; 84pp; English.

XX This polypeptide, designated APP-REP 761, comprises an amyloid

CC precursor protein (APP) that has a 276-amino acid deletion of the

CC native APP and which carries a Substance P epitope markers placed

CC N-terminal to the beta-amyloid protein (BAP) domain. APP-REP 751

CC can be used in a claimed method for screening for a compound which  
 CC reduces the formation of beta-amyloid protein, determined by  
 CC measuring the amount of marker in a medium containing transfected  
 CC cells. The method is used to detect compounds which inhibit the  
 CC activity of proteolytic enzymes which cleave APP to generate BAP  
 CC fragments. Such compounds can be used in the treatment of e.g.  
 CC Alzheimer's disease. The deletion of a 276 amino acid portion of  
 CC APP distinguishes the construct from endogenously expressed APP,  
 CC and beneficially increases the resolution of APP-REP fragments  
 CC resulting from the proteolytic cleavage by secretase or other  
 CC amyloidogenic, BAP-generating cleavage events.

XX Sequence 487 AA;

Query Match 67.6%; Score 48; DB 18; Length 487;  
 Best Local Similarity 81.8%; Pred. No. 9.7;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOWFWLM 11

|||||:|

Db 362 rpkpqffgilm 372

RESULT 123

AAW26394

ID AAW26394 standard; Protein; 487 AA.

XX AC AAW26394;

XX DT 15-DEC-1997 (first entry)

XX DE Amyloid precursor protein substrate APP-REP 751.

XX KW Amyloid precursor protein; APP; beta-amyloid protein; BAP;  
 KW substrate; mutein; secretase; Alzheimer's disease; human;  
 KW APP-REP 751; pCLL621.

XX OS Chimeric Homo sapiens;

OS Chimeric synthetic.

XX FH Key Location/Qualifiers

FT Peptide 362..372

FT /label= SP

FT /note= "substance P reporter epitope"

FT Domain 389..430

FT /label= BAP

FT /note= "beta-amyloid protein"

FT Cleavage-site 404..405

FT /note= "secretase cleavage site"

FT Domain 417..440

FT /label= Transmembrane

XX US5652092-A.

PN 29-JUL-1997.

PD 01-MAY-1992; 92US-0877675.

PF 20-SEP-1993; 93US-0123659.

PR 01-MAY-1992; 92US-0877675.

XX 05-JUN-1995; 95US-0462859.

XX (AMCY ) AMERICAN CYANAMID CO.

XX Jacobsen JS, Vitek MP;

XX WPI; 1997-392937/36.

XX N-PSDB; AAT84562.

XX Screening for compounds which reduce beta-amyloid protein formation  
 PT - using cells which express a construct encoding a marker and an  
 PT amyloid precursor mutein derived from APP isoforms

XX Disclosure; Fig 8; 84pp; English.

PS This polypeptide, designated APP-REP 761, comprises an amyloid

XX precursor protein (APP) that has a 276-amino acid deletion of the

CC native APP and which carries a Substance P epitope markers placed

CC N-terminal to the beta-amyloid protein (BAP) domain. APP-REP 751

CC can be used in a claimed method for screening for a compound which

CC reduces the formation of beta-amyloid protein, determined by

CC measuring the amount of marker in a medium containing transfected

CC cells. The method is used to detect compounds which inhibit the

CC activity of proteolytic enzymes which cleave APP to generate BAP

CC fragments. Such compounds can be used in the treatment of e.g.

CC Alzheimer's disease. The deletion of a 276 amino acid portion of

CC APP distinguishes the construct from endogenously expressed APP,

CC and beneficially increases the resolution of APP-REP fragments

CC resulting from the proteolytic cleavage by secretase or other

CC amyloidogenic, BAP-generating cleavage events.

XX Sequence 487 AA;

Query Match 67.6%; Score 48; DB 18; Length 487;

Best Local Similarity 81.8%; Pred. No. 9.7;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOWFWLM 11

DB 362 rpkpqffglm 372

|||||:| |

RESULT 124

AAW4745

ID AAW4745 standard; Protein; 487 AA.

XX

AC AAW4745;

XX

DT 01-JUN-1998 (first entry)

XX

DE APP-REP 751 protein from pCLL621.

XX

KW Amyloid precursor protein; APP; APP 751 isoform; deletion; substrate P;

KW epitope; Met-enkephalin; detection; secretase; beta-amyloid protein; BAP;

KW Alzheimer's disease; cleavage.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN US5693478-A.

XX

PD 02-DEC-1997.

XX

PF 05-JUN-1995; 95US-0464247.

XX

PR 20-SEP-1993; 93US-0123659.

PR 01-MAY-1992; 92US-0877675.

PR 05-JUN-1995; 95US-0464247.

XX

PA (AMCY ) AMERICAN CYANAMID CO.

XX

PI Jacobsen JS, Vitek MP;

XX

DR WPI: 1998-031744/03.

DR N-PSDB; AAV05850.

XX

XX Amyloid precursor muten reporter molecule assay containing antibody

PT recognised marker - used to study pathways associated with

PT Alzheimer's disease

XX

PS Disclosure; Fig 8; 84pp; English.

XX

XX This is the amino acid sequence of a novel amyloid precursor protein

CC (APP) designated APP-REP 751, contained in construct pCLL621. The

CC sequence comprises a mutant version of the APP 751 isoform of human APP

CC which contains a deletion of 276 amino acids from the central region.

CC The deleted region is replaced by a substrate P reporter epitope

CC sequence (RPKPQOWFWLM). In contrast to the APP-REP 751 encoded by the

CC construct pCLL602 (AAW4744), this sequence does not contain a

CC Met-enkephalin reporter epitope (YGFW) fused at the C-terminus of the

CC coding sequence. The shorter protein is generated for ease of detection

CC based on size difference with the wild type APP protein and also by

CC detection of the reporter epitopes. The mutant protein can be used in a

CC method to study secretase and beta-amyloid protein (BAP)-generating

CC pathways associated with Alzheimer's disease by studying proteolytic

CC cleavage of the reporter polypeptides.

XX Sequence 487 AA;

Query Match 67.6%; Score 48; DB 19; Length 487;

Best Local Similarity 81.8%; Pred. No. 9.7;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOWFWLM 11

DB 362 rpkpqffglm 372

|||||:| |

RESULT 125

AAW42979

ID AAW42979 standard; Protein; 487 AA.

XX

XX AAW42979;

XX

DT 01-MAY-1998 (first entry)

XX

DE Amyloid precursor protein mutant APP-ARP 751.

XX

KW Beta-amyloid peptide; BAP; extracellular BAP plaque;

KW cerebrovascular deposit; Alzheimers disease; Downs syndrome;

KW amyloid precursor protein; APP; secretase; BAP aggregation;

KW abnormal proteolytic cleavage.

XX

OS Synthetic.

OS Homo sapiens.

XX

PN US5703209-A.

XX

PD 30-DEC-1997.

XX

PF 05-JUN-1995; 95US-0464248.

XX

PR 20-SEP-1993; 93US-0123659.

PR 01-MAY-1992; 92US-0877675.

XX

PA (AMCY ) AMERICAN CYANAMID CO.

XX

PI Jacobsen JS, Vitek MP;

XX

DR WPI: 1998-076482/07.

DR N-PSDB; AAV04866.

XX

PT Amyloid precursor protein fusion polypeptides - comprising APP

PT fragment and marker, useful for research and drug screening

XX

PS Disclosure; Fig 8A-Q; 84pp; English.

XX

CC The present sequence represents an amyloid precursor protein (APP),

CC which has a deletion of 276 amino acids to within 15 amino acids of the

CC beta-amyloid peptide (BAP) domain. The protein also contains the Abnormal

CC accumulation of extracellular BAP in plaques and cerebrovascular deposits

CC is characteristic in brains of individuals suffering from Alzheimers

CC disease and Downs syndrome. BAP is a poorly soluble, self-aggregating

CC protein which is derived from a larger amyloid precursor protein (APP).

CC APP is expressed as an integral membrane protein, and is cleaved by

CC secretase, between BAP 16Lys and 17Leu. Cleavage at this site precludes

CC amyloidogenesis and results in the release of the amino-terminal APP  
CC fragment. Three major isoforms of APP exist: APP-695, APP-751 and  
CC APP-770. These isoforms are derived by alternative splicing. APP-RP 751  
CC is constructed by ligating restriction fragments representing N- and  
CC C-terminal APP-751 cDNA and substrate P reporter epitope sequences.  
CC APP can be used as a substrate for studying abnormal proteolytic cleavage  
CC which results in the release of BAP, and also to screen for drugs that  
CC will inhibit such cleavage.

XX SQ Sequence 487 AA;

Query Match 67.6%; Score 48; DB 19; Length 487;

Best Local Similarity 81.8%; Pred. No. 9.7;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWFWM 11  
|||||:|  
Db 362 rpqpqffglm 372

RESULT 126

AAW26509  
ID AAW26509 standard; Protein; 492 AA.

XX AC AAW26509;

XX DT 20-JUN-1994 (first entry)

XX DE APP-REP 751 amyloid precursor protein/reporter protein.  
XX KW Amyloid precursor protein; APP; beta amyloid protein; BAP;  
XX detection; Alzheimer's disease; Down's syndrome.

XX OS Homo sapiens.

XX PN AU9338358-A.

XX PD 04-NOV-1993.

XX PF 03-MAY-1993; 93AU-0038358.

XX PR 01-MAY-1992; 92US-0877675.

XX PA (AMCY ) AMERICAN CYANAMID CO.

XX PI Jacobsen JS, Vitek MP;

XX DR WPI; 1993-406194/51.

XX DR N-PSDB; AAQ54257.

XX PT New mutant forms of amyloid precursor protein - for detecting  
XX cpds. that modify activity of enzymes involved in precursor  
XX cleavage, also new nucleic acid encoding them  
XX PS Claim 5; Figure 7; 66pp; English.  
XX CC This mutant form of amyloid precursor protein comprises from the 5'  
XX to the 3' end a sequence encoding a marker and either (1) a  
XX sequence encoding the N-terminus of an amyloid precursor protein  
XX (APP) up to, but not including, the nucleotides encoding the beta  
XX amyloid protein (BAP) domain or (2) the BAP domain. Recombinant  
XX polypeptides generated from this proteins coding sequence can be  
XX used to detect drugs or compounds that inhibit/augment the  
XX activity of proteolytic enzymes which cleave APP to generate BAP  
XX fragments (deposition of which occurs in patients with Alzheimers  
XX disease and Down's syndrome).

XX SQ Sequence 492 AA;

XX Query Match 67.6%; Score 48; DB 14; Length 492;  
XX Best Local Similarity 81.8%; Pred. No. 9.8;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RPKPOQWFWM 11  
|||||:|  
Db 362 rpqpqffglm 372

RESULT 127

AAW26509  
ID AAW26509 standard; Protein; 492 AA.

XX AC AAW26509;

XX DT 06-JAN-1998 (first entry)

XX DE Amyloid precursor protein substrate APP-REP 751.

XX KW Amyloid precursor protein; APP; beta-amyloid protein; BAP;  
XX substrate; mutin; secretase; Alzheimer's disease; human;  
XX APP-REP 751; pCLL602.

XX OS Chimeric Homo sapiens.

XX OS Chimeric synthetic.

XX PH Key Location/Qualifiers

XX FT Peptide 362..372

XX FT /label= SP

XX FT /note= "substance P reporter epitope"

XX FT Domain 389..430

XX FT /label= BAP

XX FT /note= "beta-amyloid protein"

XX FT Cleavage-site 404..405

XX FT /note= "secretase cleavage site"

XX FT Domain 417..440

XX FT /label= Transmembrane

XX FT Peptide 488..492

XX FT /label= ME

XX FT /note= "Met-enkephalin reporter epitope"

XX PN US5656477-A.

XX PD 12-AUG-1997.

XX PF 01-MAY-1992; 92US-0877675.

XX PR 20-SEP-1993; 93US-0123659.

XX PR 01-MAY-1992; 92US-0877675.

XX PA (AMCY ) AMERICAN CYANAMID CO.

XX PI Jacobsen JS, Vitek MP;

XX DR WPI; 1997-414594/38.

XX DR P-PSDB; AAT87083.

XX PT Nucleic acid encoding amyloid precursor mutin(s) - comprising  
XX reporter gene and coding sequence, for identifying compounds which  
XX modify the activity of proteolytic enzymes which cleave APP

XX PS Disclosure; Fig 7; 84pp; English.

XX CC This polypeptide, designated APP-REP 761, comprises an amyloid  
XX precursor protein (APP) that has a 276-amino acid deletion of the  
XX native APP and which carries Substance P and Met-enkephalin epitope  
XX markers placed, respectively, on the N-terminal and C-terminal  
XX sites of the beta-amyloid protein (BAP) domain. APP-REP 751 can  
XX be used in a claimed method for screening for a compound which  
XX reduces the formation of beta-amyloid protein, determined by  
XX measuring the amount of marker in a medium containing transfected  
XX cells. The method is used to detect compounds which inhibit the  
XX activity of proteolytic enzymes which cleave APP to generate BAP  
XX fragments. Such compounds can be used in the treatment of e.g.  
XX Alzheimer's disease. The deletion of a 276 amino acid portion of

CC APP distinguishes the construct from endogenously expressed APP,  
CC and beneficially increases the resolution of APP-REP fragments  
CC resulting from the proteolytic cleavage by secretase or other  
CC amyloidogenic, BAP-generating cleavage events.

XX Sequence 492 AA;

Query Match 67.6%; Score 48; DB 18; Length 492;  
Best Local Similarity 81.8%; Pred. No. 9.8;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Caps 0;

QY 1 RPKPQOWFLM 11  
|||||:| |  
Db 362 rpkpqffglm 372

RESULT 128

AAW26393  
ID AAW26393 standard; Protein; 492 AA.

XX AC AAW26393;

XX 15-DEC-1997 (first entry)

DE Amyloid precursor protein substrate APP-REP 751.

KW Amyloid precursor protein; APP; beta-amyloid protein; BAP;  
KW substrate; muten; secretase; Alzheimer's disease; human;  
KW APP-REP 751; pCLL602.

XX Chimeric Homo sapiens;  
OS Chimeric synthetic.

XX Key Location/Qualifiers

FT Peptide 362..372  
FT /label= SP  
FT /note= "substance P reporter epitope"

FT Domain 389..430

FT /label= BAP  
FT /note= "beta-amyloid protein"

FT Cleavage-site 404..405

FT /note= "secretase cleavage site"

FT Domain 417..440

FT /label= Transmembrane

FT Peptide 488..492

FT /label= ME  
FT /note= "Met-enkephalin reporter epitope"

XX US5652092-A.

XX 29-JUL-1997.

XX 01-MAY-1992; 92US-0877675.

XX 20-SEP-1993; 93US-0123659.

XX 01-MAY-1992; 92US-0877675.

XX 05-JUN-1995; 95US-0462859.

XX (AMCY ) AMERICAN CYANAMID CO.

XX Jacobsen JS, Vitek MP;

XX WPI; 1997-392937/36.

XX N-PSDB; AAT84561.

XX Screening for compounds which reduce beta-amyloid protein formation  
PT - using cells which express a construct encoding a marker and an  
PT amyloid precursor muten derived from APP isoforms

XX Disclosure; Fig 7; 84pp; English.

XX This polypeptide, designated APP-REP 761, comprises an amyloid

CC precursor protein (APP) that has a 276-amino acid deletion of the  
CC native APP and which carries Substance P and Met-enkephalin epitope  
CC markers placed, respectively, on the N-terminal and C-terminal  
CC sites of the beta-amyloid protein (BAP) domain. APP-REP 751 can  
CC be used in a claimed method for screening for a compound which  
CC reduces the formation of beta-amyloid protein, determined by  
CC measuring the amount of marker in a medium containing transfected  
CC cells. The method is used to detect compounds which inhibit the  
CC activity of proteolytic enzymes which cleave APP to generate BAP  
CC fragments. Such compounds can be used in the treatment of e.g. of  
CC Alzheimer's disease. The deletion of a 276 amino acid portion of  
CC APP distinguishes the construct from endogenously expressed APP,  
CC and beneficially increases the resolution of APP-REP fragments  
CC resulting from the proteolytic cleavage by secretase or other  
CC amyloidogenic, BAP-generating cleavage events.

XX Sequence 492 AA;

Query Match 67.6%; Score 48; DB 18; Length 492;  
Best Local Similarity 81.8%; Pred. No. 9.8;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOWFLM 11  
|||||:| |

Db 362 rpkpqffglm 372

RESULT 129

AAW44744

ID AAW44744 standard; Protein; 492 AA.

XX AC AAW44744;

XX 01-JUN-1998 (first entry)

XX APP-REP 751 protein from pCLL602.

KW Amyloid precursor protein; APP; APP 751 isoform; deletion; substrate P;  
KW epitope; Met-enkephalin; detection; secretase; beta-amyloid protein; BAP;  
KW Alzheimer's disease; cleavage.

XX Homo sapiens.

OS Synthetic.

XX US5693478-A.

XX 02-DEC-1997.

XX 05-JUN-1995; 95US-0464247.

XX 20-SEP-1993; 93US-0123659.

XX 01-MAY-1992; 92US-0877675.

XX 05-JUN-1995; 95US-0464247.

XX (AMCY ) AMERICAN CYANAMID CO.

XX Jacobsen JS, Vitek MP;

XX WPI; 1998-031744/03.

XX N-PSDB; AAV05849.

XX Amyloid precursor muten reporter molecule assay containing antibody  
PT recognised marker - used to study pathways associated with  
PT Alzheimer's disease

XX Disclosure; Fig 7; 84pp; English.

XX This is the amino acid sequence of a novel amyloid precursor protein  
CC (APP) designated APP-REP 751, contained in construct pCLL602. The  
CC sequence comprises a mutant version of the APP 751 isoform of human APP  
CC which contains a deletion of 276 amino acids from the central region.  
CC The deleted region is replaced by a substrate P reporter epitope sequence

CC (RPKPOQFFGLM) and a Met-enkephalin reporter epitope (YGGFM) is fused at  
CC the C-terminus. The shorter protein is generated for ease of detection  
CC based on size difference with the wild type APP protein and also by  
CC detection of the reporter epitopes. The mutant protein can be used in  
CC a method to study secretase and beta-amyloid protein (BAP)-generating  
CC pathways associated with Alzheimer's disease by studying proteolytic  
CC cleavage of the reporter polypeptides.  
XX  
SQ Sequence 492 AA;  
  
Query Match 67.6%; Score 48; DB 19; Length 492;  
Best Local Similarity 81.8%; Pred. No. 9.8;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RPKPOQFFGLM 11  
Db 362 rpkpqffglm 372  
|||||: ||  
  
RESULT 130  
AAW42978  
ID AAW42978 standard; Protein; 492 AA.  
AC AAW42978;  
DT 01-MAY-1998 (first entry)  
XX  
DE Amyloid precursor protein mutant APP-APP 751.  
XX  
DE Beta-amyloid peptide; BAP; extracellular BAP plaque;  
KW cerebrovascular deposit; Alzheimers disease; Downs syndrome;  
KW amyloid precursor protein; APP; secretase; BAP aggregation;  
KW abnormal proteolytic cleavage.  
XX  
OS Synthetic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Protein 1..487  
FT /note= "APP-APP 751"  
FT Peptide 488..492  
FT /note= "Met-enkephalin reporter epitope"  
XX  
PN US5703209-A.  
XX  
PD 30-DEC-1997.  
XX  
XX 05-JUN-1995; 95US-0464248.  
XX  
PR 20-SEP-1993; 93US-0123659.  
PR 01-MAY-1992; 92US-0877675.  
XX  
PA (AMCY ) AMERICAN CYANAMID CO.  
XX  
PI Jacobsen JS, Vitek MP;  
XX  
XX WPI; 1998-076482/07.  
DR N-PSDB; AAV04865.  
XX  
XX Amyloid precursor protein fusion polypeptides - comprising APP  
PT fragment and marker, useful for research and drug screening  
XX  
PS Disclosure; Fig 7A-Q; 84pp; English.  
XX  
CC The present sequence represents an amyloid precursor protein (APP),  
CC which has a deletion of 276 amino acids to within 15 amino acids of the  
CC beta-amyloid peptide (BAP) domain. The protein also contains the  
CC Met-enkephalin reporter epitope at the carboxy terminus. Abnormal  
CC accumulation of extracellular BAP in plaques and cerebrovascular deposits  
CC is characteristic in brains of individuals suffering from Alzheimers  
CC disease and Downs syndrome. BAP is a poorly soluble, self-aggregating  
CC protein which is derived from a larger amyloid precursor protein (APP).  
XX

CC APP is expressed as an integral membrane protein, and is cleaved by  
CC secretase, between BAP 16lys and 17Leu. Cleavage at this site precludes  
CC amyloidogenesis and results in the release of the amino-terminal APP  
CC fragment. Three major isoforms of APP exist: APP-695, APP-751 and  
CC APP-770. These isoforms are derived by alternative splicing. APP-RP 751  
CC is constructed by ligating restriction fragments representing N- and  
CC C-terminal APP-751 cDNA and substrate P reporter epitope sequences.  
CC APP can be used as a substrate for studying abnormal proteolytic cleavage  
CC which results in the release of BAP, and also to screen for drugs that  
XX will inhibit such cleavage.  
SQ Sequence 492 AA;  
  
Query Match 67.6%; Score 48; DB 19; Length 492;  
Best Local Similarity 81.8%; Pred. No. 9.8;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RPKPOQFFGLM 11  
Db 362 rpkpqffglm 372  
|||||: ||  
  
RESULT 131  
AAR21958  
ID AAR21958 standard; Peptide; 11 AA.  
AC AAR21958;  
DT 25-JUN-1992 (first entry)  
XX  
DE Substance P [Ala 9] or [D-Ala 9].  
XX  
KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 9  
FT /note= "either L or D form"  
XX  
PN W09202248-A.  
XX  
PD 20-FEB-1992.  
XX  
XX 29-JUL-1991; 91WO-US05323.  
XX  
XX 27-JUL-1990; 90US-0559173.  
XX  
PA (CHIL-) CHILDRENS MED CENT.  
XX  
PI Yankner BA;  
XX  
XX WPI; 1992-079804/10.  
XX  
XX Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX  
XX Claim 10; Page 21; 35pp; English.  
XX  
XX The peptide is the tachykinin agonist substance P with an Ala (D/L)  
CC residue substituted at position 9. The peptide was synthesised  
CC by standard solid phase synthesis. Neuronal accumulation of  
CC beta-amyloid may be treated by administration of tachykinin  
CC agonists. The peptide can reduce the neurotoxic effects of a beta-  
CC amyloid related polypeptide on cultured neurons. The peptide and  
CC its analogues are useful for controlling diseases characterised by  
CC beta amyloid accumulation in the brain such as Alzheimer's disease  
CC and Down's syndrome.  
CC See also AAR21932-75.  
XX

SQ Sequence 11 AA;

Query Match 66.2%; Score 47; DB 13; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.36;  
 Matches 9; Conservative 1; Mismatches 0; Indels 1; Gaps 0;

QY 1 RPKPQQWFWM 11  
 |||||:| |  
 Db 1 rpkpqgffalm 11

RESULT 132  
 AAW92674  
 ID AAW92674 standard; peptide; 11 AA.  
 XX  
 AC AAW92674;  
 XX  
 DT 30-APR-1999 (first entry)  
 XX  
 DE Human tachykinin agonist beta-amyloid peptide fragment #20.  
 XX  
 DE Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
 KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
 KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5876948-A.  
 XX  
 PD 02-MAR-1999.  
 XX  
 PF 27-JUL-1991; 91US-0737371.  
 XX  
 PR 29-JUL-1991; 91US-0737371.  
 PR 27-JUL-1990; 90US-0559173.  
 XX  
 PA (CHIL-) CHILDRENS MEDICAL CENT.  
 XX  
 PI Yankner BA;  
 XX  
 PD WPI; 1999-189630/16.  
 XX  
 PF Screening for neurotoxin inhibitors - by testing compounds for their  
 PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
 XX  
 PS Disclosure; Column 19-20; 28pp; English.  
 XX  
 CC This invention describes a method for screening compounds for inhibiting  
 CC a neurotoxin. The method involves incubating tachykinin agonists with  
 CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
 CC used for identifying compounds for treating diseases characterised by an  
 CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
 CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
 CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
 CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
 CC beta-amyloid peptide fragments.  
 XX  
 SQ Sequence 11 AA;

Query Match 66.2%; Score 47; DB 20; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.36;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
 |||||:| |  
 Db 1 rpkpqgffalm 11

RESULT 133  
 AAW92675  
 ID AAW92675 standard; peptide; 11 AA.  
 XX  
 AC AAW92675;  
 XX  
 DT 25-JUN-1992 (first entry)  
 XX  
 DE Substance P [Pro 9] or [D-Pro 9].  
 XX  
 KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
 KW syndrome; hereditary cerebral haemorrhage.  
 XX  
 OS Synthetic.

XX AAW92675;  
 AC  
 DT 30-APR-1999 (first entry)  
 XX  
 DE Human tachykinin agonist beta-amyloid peptide fragment #21.  
 XX  
 KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
 KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
 KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 9  
 FT /note= "D-form residue"  
 XX  
 PN US5876948-A.  
 XX  
 PD 02-MAR-1999.  
 XX  
 PF 27-JUL-1991; 91US-0737371.  
 XX  
 PR 29-JUL-1991; 91US-0737371.  
 PR 27-JUL-1990; 90US-0559173.  
 XX  
 PA (CHIL-) CHILDRENS MEDICAL CENT.  
 XX  
 PI Yankner BA;  
 XX  
 PD WPI; 1999-189630/16.  
 XX  
 PF Screening for neurotoxin inhibitors - by testing compounds for their  
 PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
 XX  
 PS Disclosure; Column 19-20; 28pp; English.  
 XX  
 CC This invention describes a method for screening compounds for inhibiting  
 CC a neurotoxin. The method involves incubating tachykinin agonists with  
 CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
 CC used for identifying compounds for treating diseases characterised by an  
 CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
 CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
 CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
 CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
 CC beta-amyloid peptide fragments.  
 XX  
 SQ Sequence 11 AA;

Query Match 66.2%; Score 47; DB 20; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.36;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
 |||||:| |  
 Db 1 rpkpqgffalm 11

RESULT 134  
 AAR21935  
 ID AAR21935 standard; Protein; 11 AA.  
 XX  
 AC AAR21935;  
 XX  
 DT 25-JUN-1992 (first entry)  
 XX  
 DE Substance P [Pro 9] or [D-Pro 9].  
 XX  
 KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
 KW syndrome; hereditary cerebral haemorrhage.  
 XX  
 OS Synthetic.

| Key                                                                                                                                                                                                                                                                                                                                                                                                                                          | Location/Qualifiers         |
|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------|
| Modified-site                                                                                                                                                                                                                                                                                                                                                                                                                                | 9                           |
|                                                                                                                                                                                                                                                                                                                                                                                                                                              | /note= "either L or D form" |
| WO9202248-A.                                                                                                                                                                                                                                                                                                                                                                                                                                 |                             |
| 20-FEB-1992.                                                                                                                                                                                                                                                                                                                                                                                                                                 |                             |
| 29-JUL-1991;                                                                                                                                                                                                                                                                                                                                                                                                                                 | 91WO-US05323.               |
| 27-JUL-1990;                                                                                                                                                                                                                                                                                                                                                                                                                                 | 90US-0559173.               |
| (CHIL-) CHILDRENS MED CENT.                                                                                                                                                                                                                                                                                                                                                                                                                  |                             |
| Yankner BA;                                                                                                                                                                                                                                                                                                                                                                                                                                  |                             |
| WPI; 1992-079804/10.                                                                                                                                                                                                                                                                                                                                                                                                                         |                             |
| Treatment of neuronal accumulation of beta-tachykinin agonists e.g. substance P, physalgin, and bombesin, in the brains of mice with Down's disease, Downs                                                                                                                                                                                                                                                                                   |                             |
| Claim 10; Page 21; 35pp; English.                                                                                                                                                                                                                                                                                                                                                                                                            |                             |
| The peptide is the tachykinin agonist substituted at position 9. The peptide residue substituted at position 9. The peptide by standard solid phase synthesis. Neuronal beta-amyloid may be treated by administration of a tachykinin agonist. The peptide can reduce the neuronal related amyloid peptide on cultured neurons. Its analogues are useful for controlling the beta amyloid accumulation in the brain such as Down's syndrome. |                             |
| See also AAR21932-75.                                                                                                                                                                                                                                                                                                                                                                                                                        |                             |
| Sequence                                                                                                                                                                                                                                                                                                                                                                                                                                     | 11 AA;                      |

Claim 10; Page 21; 35pp; English.

The peptide is the tachykinin agonist substance P with a Pro (D/L) residue substituted at position 9. The peptide was synthesised by standard solid phase synthesis. Neuronal accumulation of beta-amyloid may be treated by administration of tachykinin agonists. The peptide can reduce the neurotoxic effects of a beta amyloid related polypeptide on cultured neurons. The peptide and its analogues are useful for controlling diseases characterised by beta amyloid accumulation in the brain such as Alzheimer's disease and Down's syndrome.

See also AAR21932-75.

```

Query Match      64.8%; Score 46; DB 13; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.51;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 RPKPQQWFWM 11
        |||||:|
Db      1 rpqpqffplm 11

RESULT 135
AAR21943
ID AAR21943 standard; Protein; 11 AA.
XX
XX AAR21943;
XX
XX
XX 25-JUN-1992 (first entry)
XX
XX Substrate P [Met 7].
XX
XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;
XX syndrome; hereditary cerebral haemorrhage.
XX
XX Synthetic.
XX
XX WO9202248-A.
XX
XX 20-FEB-1992.
XX
XX 29-JUL-1991; 91WO-US05323.
XX
XX 27-JUL-1990; 90US-0559173.
XX
XX (CHIL-) CHILDRENS MED CENT.
XX

```

|            |                                                                           |
|------------|---------------------------------------------------------------------------|
| PI         | Yankner BA;                                                               |
| XX         |                                                                           |
| DR         | WPI; 1992-079804/10.                                                      |
| XX         |                                                                           |
| PT         | Treatment of neuronal accumulation of beta-amyloid - using                |
| FT         | tachykinin agonists e.g. substance P, physalaemin and neurokinin          |
| PT         | B, for treating Alzheimer's disease, Downs syndrome, etc.                 |
| XX         |                                                                           |
| PS         | Claim 10; Page 21; 35pp; English.                                         |
| XX         |                                                                           |
| CC         | The peptide is the tachykinin agonist substance P with a                  |
| CC         | methionine residue substituted at position 7. The peptide was             |
| CC         | synthesised by standard solid phase synthesis. Neuronal                   |
| CC         | accumulation of beta-amyloid may be treated by administration of          |
| CC         | tachykinin agonists. The peptide can reduce the neurotoxic effects        |
| CC         | of a beta-amyloid related polypeptide on cultured neurons. The            |
| CC         | peptide and its analogues are useful for controlling diseases             |
| CC         | characterised by beta amyloid accumulation in the brain such as           |
| CC         | Alzheimer's disease and Down's syndrome.                                  |
| CC         | See also AAR21932-75.                                                     |
| XX         |                                                                           |
| SQ         | Sequence 11 AA;                                                           |
|            |                                                                           |
|            | Query Match 64.8%; Score 46; DB 13; Length 11;                            |
|            | Best Local Similarity 81.8%; Pred. No. 0.51;                              |
|            | Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;                |
| Qy         | 1 RPKPQQWFGLM 11                                                          |
|            | 11                                                                        |
| Db         | 1 rpqpqmgfml 11                                                           |
|            |                                                                           |
| RESULT 136 |                                                                           |
| AAW13611   |                                                                           |
| ID         | AAW13611 standard; peptide; 11 AA.                                        |
| XX         |                                                                           |
| AC         | AAW13611;                                                                 |
| XX         |                                                                           |
| DT         | 06-JUL-1999 (first entry)                                                 |
| XX         |                                                                           |
| DE         | Spantide II, a substance P antagonist.                                    |
| XX         |                                                                           |
| KW         | Substance P; antagonist; non-photosynthetic filamentous bacterium; pain;  |
| KW         | sendide; CNS; central nervous system; respiration; allergy; inflammation; |
| KW         | gastrointestinal disorder; skin; fibrosis; collagen maturation; mucosa;   |
| KW         | cardiovascular; vasospasm; immunological disorder; urinary tract;         |
| XX         | irritation.                                                               |
| XX         |                                                                           |
| OS         | Synthetic.                                                                |
| OS         | Vitreoscilla filiformis.                                                  |
| XX         |                                                                           |
| FH         | Key Location/Qualifiers                                                   |
| FT         | Modified-site 1                                                           |
| FT         | /note= "Lys-Nic; D-form residue"                                          |
| FT         | Modified-site 3                                                           |
| FT         | /note= "3-pyridyl-alanine"                                                |
| FT         | Modified-site 5                                                           |
| FT         | /note= "dichlorophenylalanine; D-form residue"                            |
| FT         | Misc-difference 3                                                         |
| FT         | /note= "D-form residue"                                                   |
| FT         | Misc-difference 9                                                         |
| FT         | /note= "D-form residue"                                                   |
| FT         | Modified-site 11                                                          |
| FT         | /label= Nle                                                               |
| FT         | /note= "Nor-leucine; amidated C-terminus"                                 |
| XX         |                                                                           |
| PN         | EP761204-Al.                                                              |
| XX         |                                                                           |
| PD         | 12-MAR-1997.                                                              |
| XX         |                                                                           |
| XX         | 13-AUG-1996; 96EP-0401781.                                                |
| XX         |                                                                           |

PR 27-MAR-1996; 96FR-0003818.  
PR 07-SEP-1995; 95FR-0010485.  
PR 27-MAR-1996; 96FR-0003816.  
PA (OREA ) L'OREAL SA.  
XX  
XX Aubert L, Breton L, De Lacharriere O, Leclaire J;  
PI Martin R;  
PI  
XX WPI; 1997-156643/15.  
XX  
PT Use of extracts of non-photosynthetic filamentous bacteria as  
PT substance P antagonists - in cosmetic and pharmaceutical compans.  
XX  
XX Claim 31; Page 7; 32pp; French.  
XX  
XX This peptide, designated spantide II, is a substance P antagonist  
CC isolated from extracts of non-photosynthetic filamentous bacteria,  
CC especially Vitreoscilla filiformis. The antagonists can be used in  
CC compositions for treating disorders associated with overproduction or  
CC excessive secretion of substance P, such as CNS disorders, respiratory  
CC disorders, allergies, inflammation, pain, gastrointestinal disorders,  
CC skin disorders, fibrosis, collagen maturation disorders, cardiovascular  
CC disorders, vasospasm, immunological disorders and/or disorders of the  
CC urinary tract; for treating sensitive skin; for preventing and/or  
CC combating skin and/or mucosal irritation.  
XX  
SQ Sequence 11 AA;  
  
Query Match 64.8%; Score 46; DB 18; Length 11;  
Best Local Similarity 60.0%; Pred. No. 0.51;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 RPKPQQHFWL 10  
Db : | | | | |  
1 kpxpxnwfwl 10  
  
RESULT 137  
AAW92677  
ID AAW92677 standard; peptide; 11 AA.  
XX  
AC AAW92677;  
XX  
XX 30-APR-1999 (first entry)  
XX  
XX Human tachykinin agonist beta-amyloid peptide fragment #23.  
XX  
XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX  
OS Homo sapiens.  
XX  
XX US5876948-A.  
XX  
XX 02-MAR-1999.  
XX  
XX 27-JUL-1991; 91US-0737371.  
XX  
XX 29-JUL-1991; 91US-0737371.  
PR 27-JUL-1990; 90US-0559173.  
XX  
XX (CHIL-) CHILDRENS MEDICAL CENT.  
XX  
XX Yankner BA;  
XX  
XX WPI; 1999-189630/16.  
XX  
XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX

PS Disclosure; Column 19-20; 28pp; English.  
XX  
XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.  
XX  
SQ Sequence 11 AA;  
  
Query Match 64.8%; Score 46; DB 20; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.51;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 RPKPQQHFWLM 11  
Db : | | | | | | | | | |  
1 rpkpqffplm 11  
  
RESULT 138  
AAW92678  
ID AAW92678 standard; peptide; 11 AA.  
XX  
AC AAW92678;  
XX  
XX 30-APR-1999 (first entry)  
XX  
XX Human tachykinin agonist beta-amyloid peptide fragment #24.  
XX  
XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 9 /note= "D-form residue"  
FT  
XX US5876948-A.  
XX  
XX 02-MAR-1999.  
XX  
XX 27-JUL-1991; 91US-0737371.  
XX  
XX 29-JUL-1991; 91US-0737371.  
PR 27-JUL-1990; 90US-0559173.  
XX  
XX (CHIL-) CHILDRENS MEDICAL CENT.  
XX  
XX Yankner BA;  
XX  
XX WPI; 1999-189630/16.  
XX  
XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX  
XX Disclosure; Column 19-20; 28pp; English.  
XX  
XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.  
XX



XX  
SQ Sequence 11 AA;

Query Match 64.8%; Score 46; DB 20; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.51;  
Matches 9; Conservative 1; Mismatches 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
|||||:| |  
Db 1 rpkpqgfplm 11

## RESULT 139

AAW92671  
ID AAW92671 standard; peptide; 11 AA.

XX  
AC AAW92671;

XX  
DT 30-APR-1999 (first entry)

XX Human tachykinin agonist beta-amyloid peptide fragment #17.

XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.

XX Homo sapiens.

XX US5876948-A.

XX 02-MAR-1999.

XX 27-JUL-1991; 91US-0737371.

XX 29-JUL-1991; 91US-0737371.

XX 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MEDICAL CENT.

XX Yankner BA;

XX WPI; 1999-189630/16.

XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells

XX Disclosure; Column 17-18; 28pp; English.

XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.

XX Sequence 11 AA;

Query Match 64.8%; Score 46; DB 20; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.51;  
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
|||||:| |  
Db 1 rpkpqgmfglm 11

## RESULT 140

AAW91415

ID AAB91415 standard; Peptide; 11 AA.

XX  
AC AAB91415;

XX 22-JUN-2001 (first entry)

XX Tachykinins peptide SEQ ID NO:591.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.

XX Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US13576.

XX 17-MAY-1999; 99US-0134406.

XX 10-SEP-1999; 99US-0153406.

XX 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity

XX Disclosure; Page 393; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.

XX Sequence 11 AA;

Query Match 64.8%; Score 46; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.51;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
|||||:| |  
Db 1 rpkpqgfplm 11

## RESULT 141

AAW91429

ID AAB91429 standard; Peptide; 11 AA.

XX  
AC AAB91429;

XX

DT 22-JUN-2001 (first entry)  
XX Tachykinins peptide SEQ ID NO:605.  
XX  
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
XX Homo sapiens.  
OS Synthetic.  
XX  
XX WO200069900-A2.  
PN  
XX 23-NOV-2000.  
PD  
XX 17-MAY-2000; 2000WO-US13576.  
PF  
XX 17-MAY-1999; 99US-0134406.  
PR  
XX 10-SEP-1999; 99US-0153406.  
PR  
XX 15-OCT-1999; 99US-0159783.  
XX  
XX (CONJ-) CONJUCHEM INC.  
PA  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
PI  
XX WPI; 2001-112059/12.  
DR  
XX  
XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
XX Disclosure; Page 397; 733pp; English.  
PS  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
XX Sequence 11 AA;  
SQ  
Query Match 64.8%; Score 46; DB 22; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.51;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQOHFWLM 11  
Db 1 rpkpqgfplm 11  
RESULT 142  
AAW50973  
ID AAW50973 standard; peptide; 8 AA.  
XX  
XX AAW50973;  
AC  
XX 31-JUL-1998 (first entry)  
DT  
XX Substance P analogue residues 4-11, [D-Pro4,D-Trp7,9,10,Phe11].  
DE  
XX

KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;  
KW Substance P; cancer; inhibition; growth hormone releasing factor;  
KW spantide.  
XX  
XX Synthetic.  
OS  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 1 /note= "D-form residue"  
FT  
FT Misc-difference 4 /note= "D-form residue"  
FT  
FT Misc-difference 6 /note= "D-form residue"  
FT  
FT Misc-difference 7 /note= "D-form residue"  
FT  
FT Misc-difference 8 /note= "D-form residue"  
FT  
FT Modified-site /note= "C-terminal amide"  
XX  
XX EP835662-A2.  
PN  
XX 15-APR-1998.  
XX  
XX 11-DEC-1996; 96EP-0309012.  
PF  
XX 08-OCT-1996; 96US-0727679.  
PR  
XX 16-AUG-1996; 96IN-0001822.  
PR  
XX (NAIW-) NAT INST IMMUNOLOGY.  
PA  
XX Jaggi M, Mukherjee R;  
XX  
XX WPI; 1998-208959/19.  
DR  
XX  
XX Composition containing analogues of vasoactive intestinal peptide,  
PT somatostatin - bombesin and substance P, for treatment of tumours  
PT and for inhibiting over-expression of these peptide(s)  
XX  
XX Disclosure; Page 13; 49pp; English.  
XX  
XX The invention relates to a new composition which comprises: (i) the  
CC somatostatin analogue SOM2 AGCKNFDWKPTSGC (3-14 disulphide bridge),  
CC and (ii) at least 4 of the peptides: antagonist of vasoactive  
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP  
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin  
CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are  
CC more general compositions containing peptide analogues of somatostatin,  
CC VIP, bombesin and substance P. The compositions are used in human or  
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour  
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,  
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,  
CC breast, kidney or particularly rectum and colon, and (b) to prevent,  
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
CC cells express receptors for VIP, somatostatin, bombesin and/or substance  
CC P. The present sequence represents a substance P analogue.  
XX  
XX Sequence 8 AA;  
SQ  
Query Match 63.4%; Score 45; DB 19; Length 8;  
Best Local Similarity 100.0%; Pred. No. 4.3e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4 PQQWF 9  
Db 1 pqqwfw 6  
RESULT 143  
AAW50975  
ID AAW50975 standard; peptide; 8 AA.  
XX  
XX AAW50975;  
AC  
XX

```
DT XX 31-JUL-1998 (first entry)
DE XX Substance P analogue residues 4-11, [D-Pro4,D-Trp7,9,Nle11].
XX Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;
KW Substance P; cancer; inhibition; growth hormone releasing factor;
KW spantide.
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH Key Misc-difference 1 /note= "D-form residue"
FT FT
FT Misc-difference 4 /note= "D-form residue"
FT FT
FT Misc-difference 6 /note= "D-form residue"
FT FT
FT Modified-site 8 /label= Nle
FT FT /note= "C-terminal amide"
FT FT
XX EP835662-A2.
PN
XX
XX 15-APR-1998.
PD
XX
XX 11-DEC-1996; 96EP-0309012.
PF
XX
XX 08-OCT-1996; 96US-0727679.
PR
XX 16-AUG-1996; 96IN-0001822.
PR
XX (NAIM-) NAT INST IMMUNOLOGY.
PA
XX
XX Jaggi M, Mukherjee R;
PI
XX WPI; 1998-208959/19.
DR
XX
XX Composition containing analogues of vasoactive intestinal peptide,
PT somatostatin - bombesin and substance P, for treatment of tumours
PT and for inhibiting over-expression of these peptide(s)
XX
XX Disclosure; Page 13; 49pp; English.
XX
XX The invention relates to a new composition which comprises: (i) the
CC somatostatin analogue SOM2 AGCKNFEdWKTpNSdC (3-14 disulphide bridge),
CC and (ii) at least 4 of the peptides: antagonist of vasoactive
CC intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP
CC receptor antagonist (VIP3); somatostatin analogue (SOM1); bombesin
CC antagonist (BOM1) and substance P antagonist (SP1). Also claimed are
CC more general compositions containing peptide analogues of somatostatin,
CC VIP, bombesin and substance P. The compositions are used in human or
CC veterinary medicine: (a) to kill (or inhibit multiplication of) tumour
CC or cancer cells, particularly for treatment of leukaemia, lymphoma,
CC adenocarcinoma of stomach, pancreas or prostate, or cancer of lung,
CC breast, kidney or particularly rectum and colon, and (b) to prevent,
CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer
CC cells express receptors for VIP, somatostatin, bombesin and/or substance
CC P. The present sequence represents a substance P analogue.
XX
XX Sequence 8 AA;

Query Match 63.4%; Score 45; DB 19; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.3e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PQQWF 9
Db 1 pqqwf 6

RESULT 144
AAR21937
ID AAR21937 standard; Protein; 11 AA.
```

```
XX AAR21937;
AC
XX 25-JUN-1992 (first entry)
DT
XX Substance P or (7-11) [Norleucine 11].
DE
XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;
KW syndrome; hereditary cerebral haemorrhage.
KW
XX Synthetic.
OS
XX Key Location/Qualifiers
FH Key Misc-difference 11 /note= OTHER - Nle"
FT FT /label=
FT FT /note= "OTHER - Nle"
FT FT
PN WO9202248-A.
XX
XX 20-FEB-1992.
PD
XX 29-JUL-1991; 91WO-US05323.
PF
XX 27-JUL-1990; 90US-0559173.
PR
XX (CHIL-) CHILDRENS MED CENT.
PA
XX Yankner BA;
PI
XX WPI; 1992-079804/10.
DR
XX Treatment of neuronal accumulation of beta-amyloid - using
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin
PT B, for treating Alzheimer's disease, Downs syndrome, etc.
XX
XX Claim 10; Page 21; 35pp; English.
XX
XX The peptide is the tachykinin agonist substance P with a Norleucine
CC residue substituted at position 11. The peptide was synthesised
CC by standard solid phase synthesis. An N-terminal deleted peptide
CC (7-11) with the same substitution was also synthesised. Neuronal
CC accumulation of beta-amyloid may be treated by administration of
CC tachykinin agonists. The peptides can reduce the neurotoxic effects
CC of a beta-amyloid related polypeptide on cultured neurons. The
CC peptide and its analogues are useful for controlling diseases
CC characterised by beta amyloid accumulation in the brain such as
CC Alzheimer's disease and Down's syndrome.
CC See also AAR21932-75.
XX
XX Sequence 11 AA;

Query Match 63.4%; Score 45; DB 13; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.72;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWF 11
Db 1 rpkpqffgll 11

RESULT 145
AAR21951
ID AAR21951 standard; Peptide; 11 AA.
XX
XX AAR21951;
AC
XX 25-JUN-1992 (first entry)
DT
XX Substance P [Glu 3].
DE
XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;
KW syndrome; hereditary cerebral haemorrhage.
KW
```

XX OS Synthetic.  
 XX PN WO9202248-A.  
 XX PD 20-FEB-1992.  
 XX PF 29-JUL-1991; 91WO-US05323.  
 XX PR 27-JUL-1990; 90US-0559173.  
 XX PA (CHIL-) CHILDRENS MED CENT.  
 XX PI Yankner BA;  
 XX DR WPI; 1992-079804/10.  
 XX PT Treatment of neuronal accumulation of beta-amyloid - using  
 PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
 PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
 XX PS Claim 10; Page 21; 35pp; English.  
 XX CC The peptide is the tachykinin agonist substance P with a glutamic  
 CC acid residue substituted at position 5. The peptide was  
 CC synthesised by standard solid phase synthesis. Neuronal  
 CC accumulation of beta-amyloid may be treated by administration of  
 CC tachykinin agonists. The peptide can reduce the neurotoxic effects  
 CC of a beta-amyloid related polypeptide on cultured neurons. The  
 CC peptide and its analogues are useful for controlling diseases  
 CC characterised by beta amyloid accumulation in the brain such as  
 CC Alzheimer's disease and Down's syndrome.  
 XX CC See also AAR21932-75.  
 XX SQ Sequence 11 AA;  
 XX

Query Match 63.4%; Score 45; DB 13; Length 11;  
 Best Local Similarity 72.7%; Pred. No. 0.72;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
 DB 1 rpkpeqffglm 11  
 |||||:| |  
 |||||:| |

RESULT 146  
 AAR28445  
 ID AAR28445 standard; peptide; 11 AA.  
 XX AC AAR28445;  
 XX DT 22-MAR-1993 (first entry)  
 XX DE Neurokinine 1 ligand #3.  
 XX KW NK1 receptor; tumour; malignant glioma; pheochromocytoma;  
 KW paraganglia; small cell lung cancer; nerve regeneration; lymphoma;  
 KW granuloma; Crohn's disease.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 11 /note= "amidated"  
 FT  
 XX PN WO9218536-A.  
 XX PD 29-OCT-1992.  
 XX PF 22-APR-1992; 92WO-US03307.  
 XX PR 22-APR-1991; 91EP-0200955.  
 XX

XX PA (MLCW ) MALLINCKRODT MEDICAL INC.  
 XX PI Bakker WH, Hagen PM, Krenning EP, Lamberts SWJ, Visser TJ;  
 XX DR WPI; 1992-382047/46.  
 XX PT Detection and localisation of tissues with neurokinine-1 receptors  
 PT - for detecting and treating tumours having neurokinine-1  
 PT receptors e.g. malignant glioma, small cell lung cancer etc.  
 XX PS Disclosure; Page 4; 22pp; English.  
 XX CC This peptide or its Tyr0 deriv. is a preferred peptide having a  
 CC selective affinity to neurokinine-1 receptors which (when  
 CC labelled with a radioactive isotope) can be used in imaging methods.  
 CC A generic formula for preferred peptides is AAR28441. Such peptides  
 CC are thus useful in diagnosis and treatment of conditions that are  
 CC related to NK1 receptors and in visualising NK1 receptors on certain  
 CC tissues. See AAR28442-R28446.  
 XX SQ Sequence 11 AA;  
 XX

Query Match 63.4%; Score 45; DB 13; Length 11;  
 Best Local Similarity 72.7%; Pred. No. 0.72;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
 DB 1 rpkpqdfyglm 11  
 |||||:| |  
 |||||:| |

RESULT 147  
 AAR42649  
 ID AAR42649 standard; peptide; 11 AA.  
 XX AC AAR42649;  
 XX DT 19-APR-1994 (first entry)  
 XX DE Neurokinin 1 receptor affinity-contg. peptide.  
 XX KW Neurokinin 1; somatostatin; receptor; cytokine; growth factor;  
 KW hormone; intra-operative; tumour; low energy gamma photon;  
 KW radionuclide.  
 XX OS Synthetic.  
 XX FH Key Location/Qualifiers  
 FT Modified-site 11 /note= "the C-terminal is amidated"  
 FT  
 XX PN WO9318797-A.  
 XX PD 30-SEP-1993.  
 XX PF 24-MAR-1993; 93WO-US02772.  
 XX PR 25-MAR-1992; 92EP-0200848.  
 XX PA (MLCW ) MALLINCKRODT MEDICAL INC.  
 XX PI Doedens BJ, Ensing GJ, Panek KJ;  
 XX DR WPI; 1993-320461/40.  
 XX PT Intra-operatively detecting and locating tumour tissues - using  
 PT specific peptide(s) labelled with low energy gamma photon  
 PT emitting radionuclide  
 XX PS Disclosure; Page 5; 31pp; English.  
 XX

CC The method of intraoperatively detecting and locating tumoral  
 CC tissues makes use of peptides having selective neurokinin 1  
 CC receptor affinity (AAR42644: generic formula; AAR42646-R42650:  
 CC specific examples), peptides having selective somatostatin  
 CC receptor affinity (AAR42645: generic formula; AAR42651-R42660:  
 CC specific examples), and peptides selected from cytokines,  
 CC growth factors and hormones.

SQ Sequence 11 AA;

Query Match 63.4%; Score 45; DB 14; Length 11;  
 Best Local Similarity 72.7%; Pred. No. 0.72;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RPKPQGFWM 11  
 |||||: ||  
 Db 1 rpkpqgfyglm 11

RESULT 148

AAW09003  
 ID AAW09003 standard; peptide; 11 AA.

AC AAW09003;

DT 03-MAR-1997 (first entry)

DE Substance P analogue, acts as substance P antagonist.

KW Analogue; substance P; spantide; non-peptide bond;  
 KW competitive inhibitor; receptor; neurogenic inflammation;  
 KW rheumatoid arthritis; ulcerative colitis; eczema; Crohn's disease;  
 KW anti-proliferative agent; small cell lung carcinoma; fibroblast.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 6..7  
 /label= Gln-psi[CH2-NH]-Phe  
 /note= "Opt. non-peptide linkage"

FT Modified-site 7..8  
 /label= Phe-psi[CH2-NH]-Phe  
 /note= "Opt. non-peptide linkage"

FT Modified-site 8..9  
 /label= Phe-psi[CH2-NH]-Gly  
 /note= "Opt. non-peptide linkage"

FT Modified-site 9..10  
 /label= Gly-psi[CH2-NH]-Leu  
 /note= "Opt. non-peptide linkage"

FT Modified-site 10..11  
 /label= Leu-psi[CH2-NH]-Leu  
 /note= "Position of claimed non-peptide linkage"

FT Modified-site 11  
 /note= "Amidated C-terminal"

FT US410019-A.

XX 25-APR-1995.

XX 24-SEP-1987; 87US-0100571.

XX 30-MAR-1992; 92US-0860675.

XX 24-SEP-1987; 87US-0100571.

XX 25-MAR-1988; 88US-0173311.

XX 08-JUN-1988; 88US-0204171.

XX 16-JUN-1988; 88US-0207759.

XX 23-SEP-1988; 88US-0248771.

XX 14-OCT-1988; 88US-0257998.

XX 09-DEC-1988; 88US-0282328.

XX 02-MAR-1989; 89US-0317941.

XX 16-AUG-1989; 89US-0394727.

PA (TULA ) TULANE EDUCATIONAL FUND.

XX Coy DH, Moreau J;

XX WPI; 1995-169633/22.

DR Novel linear peptide substance P analogues - useful as substance P  
 PT antagonists, for treating neurogenic inflammation  
 XX Claim 3; Column 19; 16pp; English.

XX The sequences given in AAW09003-04 represent analogues of substance P  
 CC and spantide, respectively. These analogues comprise a non-peptide  
 CC bond between an amino acid residue of the active site, which occurs  
 CC in the C-terminal half of the peptide, and an adjacent amino acid  
 CC residue. They act as competitive inhibitors of the naturally  
 CC occurring peptide by binding to its receptor. These peptides may be  
 CC used in the treatment of diseases involving neurogenic inflammation,  
 CC e.g. rheumatoid arthritis, ulcerative colitis, eczema and Crohn's  
 CC disease. They are also useful as anti-proliferative agents, in  
 CC the treatment of small cell lung carcinoma or disorders involving the  
 CC proliferation of fibroblasts.

XX SQ Sequence 11 AA;

Query Match 63.4%; Score 45; DB 16; Length 11;  
 Best Local Similarity 72.7%; Pred. No. 0.72;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RPKPQGFWM 11  
 |||||: ||  
 Db 1 rpkpqgfyglm 11

RESULT 149

AAW33181  
 ID AAW33181 standard; peptide; 11 AA.

AC AAW33181;

DT 29-JAN-1998 (first entry)

DE Mono-DTPA-Lys1 Substance P.

KW Substance P; radiolabel; diagnostic imaging; therapy;  
 KW mono-DTPA-Lys1.

OS Synthetic.

FH Key Location/Qualifiers

FT Modified-site 1  
 /note= "DTPA-Lys"

FT Modified-site 11  
 /note= "amidated"

FT WO9640292-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US09706.

XX 07-JUN-1995; 95US-0480372.

XX (MLCW ) MALLINCKRODT MEDICAL INC.

XX Srinivasan A;

XX WPI; 1997-087027/08.

Prepn. of pure radio-labelled peptide, e.g. for diagnostic imaging -  
 PT by combining protected poly(amino:carboxylate) ligand with peptide  
 PT and forming complex with radionuclide



## RESULT 152

AAW92679 ID AAW92679 standard; peptide; 11 AA.

XX AC AAW92679;

XX DT 30-APR-1999 (first entry)

XX DE Human tachykinin agonist beta-amyloid peptide fragment #25.

XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congenital angiopathy.

XX OS Homo sapiens.

XX FN US5876948-A.

XX PD 02-MAR-1999.

XX PF 27-JUL-1991; 91US-0737371.

XX PR 29-JUL-1991; 91US-0737371.

XX PR 27-JUL-1990; 90US-0559173.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX XX Yankner BA;

XX DR WPI; 1999-189630/16.

XX PT Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells

XX PS Disclosure; Column 21-22; 28pp; English.

XX CC This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congenital angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.

XX SQ Sequence 11 AA;

Query Match 63.4%; Score 45; DB 20; Length 11;

Best Local Similarity 72.7%; Pred. No. 0.72;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11

DB 1 RPKPQQWFWM 11

## RESULT 153

AAW91402 ID AAW91402 standard; Peptide; 11 AA.

XX AC AAW91402;

XX DT 22-JUN-2001 (first entry)

XX DE Tachykinins peptide SEQ ID NO:578.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimide; maleimide group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO2000069900-A2.

XX PD 23-NOV-2000.

XX PF 17-MAY-2000; 2000WO-US13576.

XX PR 17-MAY-1999; 99US-0134406.

XX PR 10-SEP-1999; 99US-0153406.

XX PR 15-OCT-1999; 99US-0159783.

XX PA (CONJ-) CONJUCHEM INC.

XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;

XX DR WPI; 2001-112059/12.

XX PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity

XX PS Disclosure; Page 389; 733pp; English.

XX CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimide and maleimide groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAW90829 to AAW92441 represent peptides which can be used in the  
CC exemplification of the present invention.

XX SQ Sequence 11 AA;

Query Match 63.4%; Score 45; DB 22; Length 11;

Best Local Similarity 72.7%; Pred. No. 0.72;

Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11

DB 1 RPKPQQWFWM 11

## RESULT 154

AAW91409 ID AAW91409 standard; Peptide; 11 AA.

XX AC AAW91409;

XX DT 22-JUN-2001 (first entry)

XX DE Tachykinins peptide SEQ ID NO:585.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimide; maleimide group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX OS Homo sapiens.

XX OS Synthetic.

PN WO200069900-A2.  
XX  
PD  
XX  
XX 23-NOV-2000.  
XX PF 17-MAY-2000; 2000WO-US13576.  
XX  
XX 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
XX (CONJ-) CONJUCHEM INC.  
PA  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX WPI; 2001-112059/12.  
XX  
XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
XX Disclosure; Page 391; 733pp; English.  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimide and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 63.4%; Score 45; DB 22; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.72;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
Db 1 rpkpqgfyglm 11  
|||||::||  
|::||::||

RESULT 155  
AAY37217  
ID AAY37217 standard; Protein; 455 AA.  
XX  
XX AAY37217;  
XX  
XX 07-OCT-1999 (first entry)  
DT  
XX  
XX Amino acid sequence of a Chlamydia trachomatis protein.  
DE  
XX  
XX Vaccine; eye disease; conventional trachoma; nonendemic trachoma;  
KW paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;  
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;  
KW Bartholinitis; pneumonia; venereal lymphogranulomatosis.  
XX  
XX Chlamydia trachomatis.  
OS  
XX WO9928475-A2.  
PN  
XX 10-JUN-1999.  
PD  
XX

PF 27-NOV-1998; 98WO-IB01939.  
XX  
XX 04-NOV-1998; 98US-0107077.  
PR 28-NOV-1997; 97FR-0015041.  
PR 17-DEC-1997; 97FR-0016034.  
XX  
XX (GEST ) GENSET.  
XX  
XX Griffais R;  
XX  
XX WPI; 1999-371125/31.  
XX  
XX Genome sequence of Chlamydia trachomatis  
PT  
XX  
XX Disclosure; Page 981-982; 1755pp; English.  
XX  
XX AAY36754-Y37949 are encoded by open reading frames (ORFs) of the genome  
CC of Chlamydia trachomatis (see AZ01425). The polypeptides can be used as  
CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences  
CC can also be used to control growth of the microorganism. Chlamydia  
CC trachomatis is responsible for a large number of diseases; e.g. eye  
CC diseases such as conventional trachoma, nonendemic trachoma,  
CC paratrachoma, and inclusion conjunctivitis; genital diseases such as  
CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis;  
CC perihhepatitis, Bartholinitis; pneumopathy in breast feeding infants;  
CC and venereal lymphogranulomatosis. The polypeptides of the invention  
CC may be of use in treating these diseases.  
XX  
XX Sequence 455 AA;  
SQ

Query Match 63.4%; Score 45; DB 20; Length 455;  
Best Local Similarity 75.0%; Pred. No. 26;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 3 KPOQFWL 10  
Db 303 kpeqwlw 310  
||::|||  
||::|||

RESULT 156  
AAY92927  
ID AAY92927 standard; peptide; 11 AA.  
XX  
XX AAY92927;  
AC  
XX 25-OCT-2000 (first entry)  
DT  
XX  
XX Spantide II peptide.  
DE  
XX  
XX Antiinflammatory; antipruritic; hypertensive; antidiarrhetic; analgesic;  
KW bradykinin antagonist; plant extract; Ocimum; bradykinin; antagonist;  
KW inflammation mediator; substance P; calcitonin gene related peptide;  
KW nitric oxide synthase; cytokine; histamine; tumour necrosis factor alpha;  
KW sendide; cosmetic; sensitive skin; cutaneous irritation; erythema; pain;  
KW dysesthesia; pruritis; vasodilation; hypotension; inflammation;  
KW diarrhoea; allergic rhinitis; smooth muscle contraction.  
XX  
XX Unidentified.  
OS  
XX

Key Location/Qualifiers  
FH Modified-site 1 /note= "nicotiny-Lys; D-form residue"  
FT  
FT Modified-site 3 /note= "3-pyridylalanine"  
FT  
FT Modified-site 5 /note= "di-chloro-Phe; D-form residue"  
FT  
FT Misc-difference 7 /note= "D-form residue"  
FT  
FT Misc-difference 9 /note= "D-form residue"  
FT  
FT Modified-site 11 /label= Nle  
FT





CC (H-RPKPEEFFGLM-NH2) or these peptides derived from it can be used in  
CC aq. soln. or suspension to promote hair growth and regeneration.  
CC See also AAP50632 and AAP50634.

XX Sequence 10 AA;

Query Match 60.6%; Score 43; DB 6; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQQWFWM 11  
||||:|  
Db 1 pkpqffglm 10

## RESULT 159

AAR21933  
ID AAR21933 standard; Protein; 10 AA.

XX AC AAR21933;

XX DT 25-JUN-1992 (first entry)

XX DE Substance P (2-11) fragment.

XX KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
XX syndrome; hereditary cerebral haemorrhage.

XX OS Synthetic.

XX PN WO9202248-A.

XX PD 20-FEB-1992.

XX PF 29-JUL-1991; 91WO-0505323.

XX PR 27-JUL-1990; 90US-0559173.

XX PA (CHIL-) CHILDRENS MED CENT.

XX PI Yankner BA;

XX DR WPI; 1992-079804/10.

XX PT Treatment of neuronal accumulation of beta-amyloid - using  
XX tachykinin agonists e.g. substance P, physalaemin and neurokinin  
XX B, for treating Alzheimer's disease, Downs syndrome, etc.

XX PS Claim 9; Page 21; 35pp; English.

XX CC The peptide is a tachykinin agonist consisting of residues 2-11 of  
XX substance P. The peptide was synthesised by standard solid phase  
XX syntheses. Analogues of the peptide, with N-terminal deletions down  
XX to substance P (7-11) were also synthesised. Neuronal accumulation of  
XX beta-amyloid may be treated by administration of these tachykinin  
XX agonists. The peptides reduce the neurotoxic effects of a beta-  
XX amyloid related polypeptide on cultured neurons. The peptide and  
XX its analogues are useful for controlling diseases characterised by  
XX beta amyloid accumulation in the brain such as Alzheimer's disease  
XX and Down's syndrome.  
XX See also AAR21932-75.

XX Sequence 10 AA;

Query Match 60.6%; Score 43; DB 13; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQQWFWM 11  
||||:|  
Db 1 pkpqffglm 10

## RESULT 160

AAW92663

ID AAW92663 standard; peptide; 10 AA.

XX AC AAW92663;

XX DT 30-APR-1999 (first entry)

XX DE Human tachykinin agonist beta-amyloid peptide fragment #9.

XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
XX Alzheimer's disease; Down's syndrome; amyloidosis; human;  
XX hereditary cerebral haemorrhage; non-inherited congenital angiodysplasia.

XX OS Homo sapiens.

XX PN US5876948-A.

XX PD 02-MAR-1999.

XX PF 27-JUL-1991; 91US-0737371.

XX PR 29-JUL-1991; 91US-0737371.

XX PR 27-JUL-1990; 90US-0559173.

XX PA (CHIL-) CHILDRENS MEDICAL CENT.

XX PI Yankner BA;

XX DR WPI; 1999-189630/16.

XX PT Screening for neurotoxin inhibitors - by testing compounds for their  
XX effect on beta-amyloid peptide neurotoxic effect on neuronal cells

XX PS Disclosure; Column 13-14; 28pp; English.

XX CC This invention describes a method for screening compounds for inhibiting  
XX a neurotoxin. The method involves incubating tachykinin agonists with  
XX neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
XX used for identifying compounds for treating diseases characterised by an  
XX undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
XX Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
XX with amyloidosis and non-inherited congenital angiodysplasia with cerebral  
XX haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
XX beta-amyloid peptide fragments.

XX Sequence 10 AA;

Query Match 60.6%; Score 43; DB 20; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQQWFWM 11  
||||:|  
Db 1 pkpqffglm 10

## RESULT 161

AAB91423

ID AAB91423 standard; Peptide; 10 AA.

XX AC AAB91423;

XX DT 22-JUN-2001 (first entry)

XX DE Tachykinins peptide SEQ ID NO:599.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
XX blood component; modification; succinimidyl; maleimido group; amino;  
XX hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX  
DR WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
XX  
XX  
PS Disclosure; Page 395; 733pp; English.  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 10 AA;

Query Match 60.6%; Score 43; DB 22; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFNL 10  
Db 1 rpqpqqffgl 10

RESULT 162  
AAB91427  
ID AAB91427 standard; Peptide; 10 AA.  
XX  
AC AAB91427;  
XX  
XX 22-JUN-2001 (first entry)  
XX  
DE Tachykinins peptide SEQ ID NO:603.  
XX  
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX

PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX  
DR WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
XX  
XX  
PS Disclosure; Page 396; 733pp; English.  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 10 AA;

Query Match 60.6%; Score 43; DB 22; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFNL 10  
Db 1 rpqpqqffgl 10

RESULT 163  
AAB91445  
ID AAB91445 standard; Peptide; 10 AA.  
XX  
AC AAB91445;  
XX  
XX 22-JUN-2001 (first entry)  
XX  
DE Tachykinins peptide SEQ ID NO:621.  
XX  
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX

PF 17-MAY-2000; 2000WO-US13576.  
XX  
XX 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX (CONJ-) CONJUCHEM INC.  
PA  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
PI WPI; 2001-112059/12.  
XX  
XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
XX Disclosure; Page 402; 733pp; English.  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimide and maleimide groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
XX Sequence 10 AA;  
SQ

Query Match 60.6%; Score 43; DB 22; Length 10;  
Best Local Similarity 80.0%; Pred. No. 1.3;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQOMFWLM 11  
Db 1 pkpqffglm 10

RESULT 164  
AAR21945  
ID AAR21945 standard; Protein; 11 AA.  
XX  
XX AAR21945;  
XX  
XX 25-JUN-1992 (first entry)  
XX  
XX Substance P [Pro 1].  
XX  
XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX  
XX Synthetic.  
XX  
XX WO9202248-A.  
XX  
XX 20-FEB-1992.  
XX  
XX 29-JUL-1991; 91WO-US05323.  
XX  
XX 27-JUL-1990; 90US-0559173.  
XX  
XX (CHIL-) CHILDRENS MED CENT.  
XX

Query Match 60.6%; Score 43; DB 13; Length 11;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQOMFWLM 11  
Db 2 pkpqffglm 11

RESULT 165  
AAR21936  
ID AAR21936 standard; Protein; 11 AA.  
XX  
XX AAR21936;  
XX  
XX 25-JUN-1992 (first entry)  
XX  
XX Substance P or (7-11) [Ethionine 11].  
XX  
XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 11  
FT /label= OTHER  
FT /note= "OTHER = Ethionine"  
XX  
XX WO9202248-A.  
XX  
XX 20-FEB-1992.  
XX  
XX 29-JUL-1991; 91WO-US05323.  
XX  
XX 27-JUL-1990; 90US-0559173.  
XX  
XX (CHIL-) CHILDRENS MED CENT.  
XX  
XX Yankner BA;  
XX  
XX WPI; 1992-079804/10.  
XX  
XX Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX  
XX Claim 10; Page 21; 35pp; English.  
XX

CC The peptide is the tachykinin agonist substance P with an Ethionine  
CC residue substituted at position 11. The peptide was synthesised  
CC by standard solid phase synthesis. An N-terminal deleted peptide  
CC (7-11) with the same substitution was also synthesised. Neuronal  
CC accumulation of beta-amyloid may be treated by administration of  
CC tachykinin agonists. The peptides can reduce the neurotoxic effects  
CC of a beta-amyloid related polypeptide on cultured neurons. The  
CC peptide and its analogues are useful for controlling diseases  
CC characterised by beta amyloid accumulation in the brain such as  
CC Alzheimer's disease and Down's syndrome.  
CC See also AAR21932-75.  
XX  
SQ Sequence 11 AA;

Query Match 60.6%; Score 43; DB 13; Length 11;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 10  
|||||:|  
Db 1 rpkpqqffgl 10

RESULT 166  
AAR21941  
ID AAR21941 standard; Protein; 11 AA.

XX AAR21941;

XX 25-JUN-1992 (first entry)

XX Substance P [pGLU 1].

XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
XX syndrome; hereditary cerebral haemorrhage.

XX Synthetic.

XX Key Location/Qualifiers

XX Misc-difference 1

XX /label= OTHER  
XX /note= "OTHER = pyro Glu"

XX WO9202248-A.

XX 20-FEB-1992.

XX 29-JUL-1991; 91WO-US05323.

XX 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using  
XX tachykinin agonists e.g. substance P, physalaemin and neurokinin  
XX B, for treating Alzheimer's disease, Downs syndrome, etc.

XX Claim 10; Page 21; 35pp; English.

XX The peptide is the tachykinin agonist substance P with a pyro  
XX Glutamic acid residue substituted at position 1. The peptide was  
XX synthesised by standard solid phase synthesis. Neuronal  
XX accumulation of beta-amyloid may be treated by administration of  
XX tachykinin agonists. The peptide can reduce the neurotoxic effects  
XX of a beta-amyloid related polypeptide on cultured neurons. The  
XX peptide and its analogues are useful for controlling diseases  
XX characterised by beta amyloid accumulation in the brain such as  
XX Alzheimer's disease and Down's syndrome.

CC See also AAR21932-75.  
XX Sequence 11 AA;  
SQ

Query Match 60.6%; Score 43; DB 13; Length 11;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQQWFWM 11  
|||||:|  
Db 2 pkpqqffglm 11

RESULT 167

AAR21944

ID AAR21944 standard; Protein; 11 AA.

XX AAR21944;

XX 25-JUN-1992 (first entry)

XX Substance P [Pro 11].

XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
XX syndrome; hereditary cerebral haemorrhage.

XX Synthetic.

XX WO9202248-A.

XX 20-FEB-1992.

XX 29-JUL-1991; 91WO-US05323.

XX 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using  
XX tachykinin agonists e.g. substance P, physalaemin and neurokinin  
XX B, for treating Alzheimer's disease, Downs syndrome, etc.

XX Claim 10; Page 21; 35pp; English.

XX The peptide is the tachykinin agonist substance P with a Proline  
XX residue substituted at position 11. The peptide was  
XX synthesised by standard solid phase synthesis. Neuronal  
XX accumulation of beta-amyloid may be treated by administration of  
XX tachykinin agonists. The peptide can reduce the neurotoxic effects  
XX of a beta-amyloid related polypeptide on cultured neurons. The  
XX peptide and its analogues are useful for controlling diseases  
XX characterised by beta amyloid accumulation in the brain such as  
XX Alzheimer's disease and Down's syndrome.  
XX See also AAR21932-75.

XX Sequence 11 AA;

Query Match 60.6%; Score 43; DB 13; Length 11;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 10  
|||||:|  
Db 1 rpkpqqffgl 10

RESULT 168

AAW92709  
 ID AAW92709 standard; peptide; 11 AA.  
 AC AAW92709;  
 XX  
 DT 30-APR-1999 (first entry)  
 XX  
 DE Human tachykinin agonist beta-amyloid peptide fragment #55.  
 KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
 KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
 KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 9  
 FT /label= Megly  
 FT /note= "N-methyl-glycine (sarcosine)"  
 FT Modified-site 11  
 FT /note= "Residue is Met(O2)"  
 FT  
 XX  
 PN US5876948-A.  
 XX  
 PD 02-MAR-1999.  
 XX  
 PF 27-JUL-1991; 91US-0737371.  
 XX  
 PR 29-JUL-1991; 91US-0737371.  
 PR 27-JUL-1990; 90US-0559173.  
 XX  
 PA (CHIL-) CHILDRENS MEDICAL CENT.  
 XX  
 PI Yankner BA;  
 XX  
 DR WPI; 1999-189630/16.  
 XX  
 PT Screening for neurotoxin inhibitors - by testing compounds for their  
 effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
 XX  
 PS Disclosure; Column 35-36; 28pp; English.  
 XX  
 CC This invention describes a method for screening compounds for inhibiting  
 a neurotoxin. The method involves incubating tachykinin agonists with  
 neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
 used for identifying compounds for treating diseases characterised by an  
 undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
 Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
 with amyloidosis and non-inherited congophilic angiopathy with cerebral  
 haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
 beta-amyloid peptide fragments.  
 XX  
 SQ Sequence 11 AA;

Query Match 60.6%; Score 43; DB 20; Length 11;  
 Best Local Similarity 80.0%; Pred. No. 1.4;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWL 10  
 |||||:|  
 Db 1 rpkpqqffgl 10

RESULT 169  
 AAW92717  
 ID AAW92717 standard; peptide; 11 AA.  
 AC AAW92717;  
 XX  
 DT 30-APR-1999 (first entry)  
 XX  
 DE Human tachykinin agonist beta-amyloid peptide fragment #63.

XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
 KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
 KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 11  
 FT /label= MeMet  
 FT /note= "N-methyl-methionine"  
 FT  
 XX  
 PN US5876948-A.  
 XX  
 PD 02-MAR-1999.  
 XX  
 PF 27-JUL-1991; 91US-0737371.  
 XX  
 PR 29-JUL-1991; 91US-0737371.  
 PR 27-JUL-1990; 90US-0559173.  
 XX  
 PA (CHIL-) CHILDRENS MEDICAL CENT.  
 XX  
 PI Yankner BA;  
 XX  
 DR WPI; 1999-189630/16.  
 XX  
 PT Screening for neurotoxin inhibitors - by testing compounds for their  
 effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
 XX  
 PS Disclosure; Column 37-38; 28pp; English.  
 XX  
 CC This invention describes a method for screening compounds for inhibiting  
 a neurotoxin. The method involves incubating tachykinin agonists with  
 neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
 used for identifying compounds for treating diseases characterised by an  
 undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
 Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
 with amyloidosis and non-inherited congophilic angiopathy with cerebral  
 haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
 beta-amyloid peptide fragments.  
 XX  
 SQ Sequence 11 AA;

Query Match 60.6%; Score 43; DB 20; Length 11;  
 Best Local Similarity 80.0%; Pred. No. 1.4;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWL 10  
 |||||:|  
 Db 1 rpkpqqffgl 10

RESULT 170  
 AAW92718  
 ID AAW92718 standard; peptide; 11 AA.  
 AC AAW92718;  
 XX  
 DT 30-APR-1999 (first entry)  
 XX  
 DE Human tachykinin agonist beta-amyloid peptide fragment #64.  
 XX  
 KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
 KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
 KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US5876948-A.  
 XX  
 PD 02-MAR-1999.

```

XX PF 27-JUL-1991; 91US-0737371.
XX PR 29-JUL-1991; 91US-0737371.
XX PR 27-JUL-1990; 90US-0559173.
XX PA (CHIL-) CHILDRENS MEDICAL CENT.
XX PI Yankner BA;
XX WPT; 1999-189630/16.
XX DR Screening for neurotoxin inhibitors - by testing compounds for their
XX PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells
XX PS Disclosure; Column 37-38; 28pp; English.
XX CC This invention describes a method for screening compounds for inhibiting
XX CC a neurotoxin. The method involves incubating tachykinin agonists with
XX CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be
XX CC used for identifying compounds for treating diseases characterised by an
XX CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,
XX CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage
XX CC with amyloidosis and non-inherited congophilic angiopathy with cerebral
XX CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human
XX CC beta-amyloid peptide fragments.
XX SQ Sequence 11 AA;

Query Match 60.6%; Score 43; DB 20; Length 11;
Best Local Similarity 80.0%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWL 10
Db 1 rpkpqqffgl 10

RESULT 171
AAW92667
XX ID AAW92667 standard; peptide; 11 AA.
XX AC AAW92667;
XX DT 30-APR-1999 (first entry)
XX DE Human tachykinin agonist beta-amyloid peptide fragment #13.
XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;
XX KW Alzheimer's disease; Down's syndrome; amyloidosis; human;
XX KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 11 /note= "Residue is ethionine"
XX FT
XX PN US5876948-A.
XX PD 02-MAR-1999.
XX PF 27-JUL-1991; 91US-0737371.
XX PR 29-JUL-1991; 91US-0737371.
XX PR 27-JUL-1990; 90US-0559173.
XX PA (CHIL-) CHILDRENS MEDICAL CENT.
XX PI Yankner BA;
XX WPT; 1999-189630/16.
XX DR Screening for neurotoxin inhibitors - by testing compounds for their
XX PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells
XX PS Disclosure; Column 37-38; 28pp; English.
XX CC This invention describes a method for screening compounds for inhibiting
XX CC a neurotoxin. The method involves incubating tachykinin agonists with
XX CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be
XX CC used for identifying compounds for treating diseases characterised by an
XX CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,
XX CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage
XX CC with amyloidosis and non-inherited congophilic angiopathy with cerebral
XX CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human
XX CC beta-amyloid peptide fragments.
XX SQ Sequence 11 AA;

Query Match 60.6%; Score 43; DB 20; Length 11;
Best Local Similarity 80.0%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWL 10
Db 1 rpkpqqffgl 10

RESULT 172
AAW92668
XX ID AAW92668 standard; peptide; 11 AA.
XX AC AAW92668;
XX DT 30-APR-1999 (first entry)
XX DE Human tachykinin agonist beta-amyloid peptide fragment #14.
XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;
XX KW Alzheimer's disease; Down's syndrome; amyloidosis; human;
XX KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 11 /label= Nle
XX FT
XX PN US5876948-A.
XX PD 02-MAR-1999.
XX PF 27-JUL-1991; 91US-0737371.
XX PR 29-JUL-1991; 91US-0737371.
XX PR 27-JUL-1990; 90US-0559173.
XX PA (CHIL-) CHILDRENS MEDICAL CENT.
XX PI Yankner BA;
XX WPT; 1999-189630/16.
XX DR Screening for neurotoxin inhibitors - by testing compounds for their
XX PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells
XX PS Disclosure; Column 15-16; 28pp; English.
XX CC This invention describes a method for screening compounds for inhibiting
XX CC a neurotoxin. The method involves incubating tachykinin agonists with
XX CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be
XX CC used for identifying compounds for treating diseases characterised by an
XX CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,

```

```

XX Screening for neurotoxin inhibitors - by testing compounds for their
XX PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells
XX PS Disclosure; Column 15-16; 28pp; English.
XX CC This invention describes a method for screening compounds for inhibiting
XX CC a neurotoxin. The method involves incubating tachykinin agonists with
XX CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be
XX CC used for identifying compounds for treating diseases characterised by an
XX CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,
XX CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage
XX CC with amyloidosis and non-inherited congophilic angiopathy with cerebral
XX CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human
XX CC beta-amyloid peptide fragments.
XX SQ Sequence 11 AA;

Query Match 60.6%; Score 43; DB 20; Length 11;
Best Local Similarity 80.0%; Pred. No. 1.4;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWL 10
Db 1 rpkpqqffgl 10

RESULT 172
AAW92668
XX ID AAW92668 standard; peptide; 11 AA.
XX AC AAW92668;
XX DT 30-APR-1999 (first entry)
XX DE Human tachykinin agonist beta-amyloid peptide fragment #14.
XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;
XX KW Alzheimer's disease; Down's syndrome; amyloidosis; human;
XX KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 11 /label= Nle
XX FT
XX PN US5876948-A.
XX PD 02-MAR-1999.
XX PF 27-JUL-1991; 91US-0737371.
XX PR 29-JUL-1991; 91US-0737371.
XX PR 27-JUL-1990; 90US-0559173.
XX PA (CHIL-) CHILDRENS MEDICAL CENT.
XX PI Yankner BA;
XX WPT; 1999-189630/16.
XX DR Screening for neurotoxin inhibitors - by testing compounds for their
XX PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells
XX PS Disclosure; Column 15-16; 28pp; English.
XX CC This invention describes a method for screening compounds for inhibiting
XX CC a neurotoxin. The method involves incubating tachykinin agonists with
XX CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be
XX CC used for identifying compounds for treating diseases characterised by an
XX CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,

```

CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.

XX SQ Sequence 11 AA;

Query Match 60.6%; Score 43; DB 20; Length 11;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQWFWM 10  
|||||:|  
Db 1 rpkpqffgl 10

RESULT 173  
AAW92670  
ID AAW92670 standard; peptide; 11 AA.

XX AC AAW92670;

XX DT 30-APR-1999 (first entry)

XX Human tachykinin agonist beta-amyloid peptide fragment #16.

XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "Residue is ethionine"

XX US5876948-A.

XX PD 02-MAR-1999.

XX PF 27-JUL-1991; 91US-0737371.

XX PR 29-JUL-1991; 91US-0737371.

XX PR 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MEDICAL CENT.

XX Yankner BA;

XX WPI; 1999-189630/16.

XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells

XX Disclosure; Column 17-18; 28pp; English.

XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.

XX SQ Sequence 11 AA;

Query Match 60.6%; Score 43; DB 20; Length 11;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQWFWM 11  
|||||:|  
Db 2 pkpqffgl 11

RESULT 174

AAW92672

ID AAW92672 standard; peptide; 11 AA.

XX AC AAW92672;

XX DT 30-APR-1999 (first entry)

XX Human tachykinin agonist beta-amyloid peptide fragment #18.

XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.

XX OS Homo sapiens.

XX US5876948-A.

XX PD 02-MAR-1999.

XX PF 27-JUL-1991; 91US-0737371.

XX PR 29-JUL-1991; 91US-0737371.

XX PR 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MEDICAL CENT.

XX Yankner BA;

XX WPI; 1999-189630/16.

XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells

XX Disclosure; Column 17-18; 28pp; English.

XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.

XX SQ Sequence 11 AA;

Query Match 60.6%; Score 43; DB 20; Length 11;  
Best Local Similarity 80.0%; Pred. No. 1.4;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQWFWM 11  
|||||:|  
Db 2 pkpqffgl 11

RESULT 175

AAW92672

ID AAW92672 standard; peptide; 11 AA.

XX AC AAW92672;

XX DT 20-DEC-2000 (first entry)

XX Peptide identified from a databank of polypeptides and polynucleotides.



XX Precursor peptide; polypeptide hormone; peptide identification.  
XX Unidentified.  
XX

Key Location/Qualifiers

Modified-site 1 /note= "hydrogen attached"

Modified-site 11

/note= "amidated residue"

WO2000050636-A1.

31-AUG-2000.

24-FEB-2000; 2000WO-FR00460.

25-FEB-1999; 99US-0257525.

(SCRS ) SCRS SOC CONSEILS RECH & APPL SCI.

(CNRS ) CNRS CENT NAT RECH SCI.

Camara Ferrer JJA, Thuriel C, Martinez J, Berge G, Goze C;

WPI; 2000-572101/53.

Identifying peptide with selected function, useful particularly for C-amidated hormones, by screening database for combination of nucleic acid and amino acid sequences -

Disclosure; Page 22; 40pp; French.

The specification describes a method for identifying a peptide having a particular function. The method comprises preparing a database of polynucleotides and polypeptides of unknown functions, screening the database for a combination of nucleotides or amino acids indicative of the peptide with a particular function, and identifying polynucleotides and proteins which contain the peptide. The method is used to identify precursor peptides with an amidated C-terminus, especially polypeptide hormones, for studying physiologically active substances. The present sequence represents a peptide which was identified using the method of the invention.

Sequence 11 AA;

Query Match 60.6%; Score 43; DB 21; Length 11;

Best Local Similarity 80.0%; Pred. No. 1.4;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPOQWFWM 11

||||| 11

Db 2 pkpqffglm 11

RESULT 176

AAB99350

ID AAB99350 standard; peptide; 11 AA.

AC AAB99350;

24-AUG-2001 (first entry)

Substance P tachykinin-related peptide SEQ ID NO:3.

Tachykinin-related peptide; substance P; neurokinin A; neurokinin B; physiologically active; tachykinin; drug; veterinary medicine; agrochemical.

Synthetic.

Key Location/Qualifiers

Modified-site 11

FT /note= "amidated"

XX WO200134637-A1.

XX 17-MAY-2001.

XX 07-NOV-2000; 2000WO-JP07789.

XX 08-NOV-1999; 99JP-0317535.

XX (SUNR ) SUNTORY LTD.

XX Ikeda T, Nomoto K, Minakata H;

XX WPI; 2001-329069/34.

Synthesis of new physiologically-active peptide analogs of tachykinin, useful as drugs, veterinary medicines and agrochemicals, comprises modifying C-terminal amino-acid residues -

Claim 13; Page 23; 36pp; Japanese.

The present invention describes a method for producing physiologically active substances. The method comprises converting the amino-acid residue at a specific position in a peptide into another amino-acid residue to provide activity against (in)vertebrates. Also described are: (1) converted unnatural tachykinin-related peptide with tachykinin-like physiological activity against (in)vertebrates which is an amino-acid sequence with the 5 amino-acids at C-terminal as in -Phe-AA1-Gly-AA2-Met-NH2 (1) where AA1 = Val, Ile, Phe, Tyr, His, Met, Thr, Leu, Gly or Gln and AA2 = Ser, Ala, Val, Met, Thr, Pro or Leu; and (2) drugs, veterinary medicines or agrochemicals containing the new tachykinin-related peptide as the active ingredient. The peptide analogues are for use as drugs, veterinary medicines and agrochemicals with activity on vertebrates or invertebrates. By modifying C-terminal amino-acid residues of tachykinin, the activity of tachykinin and its related derivative can therefore be changed from that against vertebrates to that on invertebrates or vice versa. The present sequence represents a specifically claimed tachykinin-related peptide from the present invention.

Sequence 11 AA;

Query Match 60.6%; Score 43; DB 22; Length 11;

Best Local Similarity 80.0%; Pred. No. 1.4;

Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 10

||||| 11

Db 1 rpkpqffgl 10

RESULT 177

AAW14777

ID AAW14777 standard; Protein; 898 AA.

AC AAW14777;

14-MAY-1997 (first entry)

Granulosis virus infectivity protein.

Infection; nuclear polyhedrosis virus; Granulosis virus genus.

Granulosis virus Xcgv alpha-4 strain.

Key Location/Qualifiers

CDS 49..2745

/\*tag= a

JP09009972-A.

PD 14-JAN-1997.  
XX  
PF 03-JUL-1995; 95JP-0167481.  
XX  
PR 03-JUL-1995; 95JP-0167481.  
XX  
XX (NORQ ) NORINSUISANSHO NOGYO KENKYU.  
PA WPI: 1997-126429/12.  
DR N-PSDB; AAT14777.  
XX  
XX Granulosis virus protein - enhances nuclear polyhedrosis virus  
PT Infectious activity  
XX  
XX Claim 4; Page 10-13; 14pp; Japanese.  
XX  
XX This sequence represents a protein which has a mol. wt. of ca. 100 KD and  
CC which enhances the infectious activity of nuclear polyhedrosis virus.  
CC This sequence is originated from a virus of Granulosis virus genus.  
CC The protein enhances infectious activity of insecticidal NPV.  
XX  
XX  
SQ Sequence 898 AA;

Query Match 60.6%; Score 43; DB 18; Length 898;  
Best Local Similarity 70.0%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQWFWM 11  
| | | | |  
DB 353 pypqiwawlm 362

## RESULT 178

AAW14285  
ID AAW14285 standard; Protein: 902 AA.

XX AC AAW14285;

DT 16-JAN-1998 (first entry)

DE H. armigera granulovirus enhancin protein.

XX Enhancin; hybridisation; granulovirus; Trichoplusia ni; insect; midgut;  
KW open reading frame; pseudaletia unipuncta; peritrophic membrane protein;  
KW insect; absorption; penetration; baculovirus; mortality; pest control.  
XX

OS Helicoverpa armigera granulovirus.

XX WO9708197-A1.

XX PD 06-MAR-1997.

XX PF 23-AUG-1996; 96WO-0513645.

XX PR 23-AUG-1996; 96US-0002743.

XX PR 24-AUG-1995; 95US-0002743.

XX PA (BOYC-) BOYCE THOMPSON INST PLANT RES.

XX PI Granados RR;

XX DR WPI: 1997-179177/16.

XX DR N-PSDB; AAT79099.

XX DNA encoding the Helicoverpa armigera granulovirus enhancin protein  
PT - used to increase infectivity of baculovirus for control of insect  
PT pests

XX Claim 1; Page 18-22; 36pp; English.

XX This is the amino acid sequence of the enhancin protein of Helicoverpa  
CC armigera granulovirus (HearGV). The protein has a calculated molecular

CC weight of 104.6 kD which is similar to the 104.2 kD of the 901 amino  
CC acid enhancin from Pseudaletia unipuncta granulovirus (PsunGV; see patent  
CC US5475090). The HearGV enhancin also has 80 and 81% amino acid identity  
CC with the Trichoplusia ni granulovirus (TnGV) and PsunGV enhancins  
CC respectively. The enhancin proteins disrupt peritrophic membrane  
CC proteins in insects so that absorption, penetration and/or uptake of  
CC baculovirus in the midgut is increased, leading to higher mortality.  
CC They are used in conjunction with a substance toxic to insects for pest  
CC control.  
XX  
XX  
SQ Sequence 902 AA;

Query Match 60.6%; Score 43; DB 18; Length 902;  
Best Local Similarity 70.0%; Pred. No. 1e+02;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQWFWM 11  
| | | | |  
DB 353 pypqiwawlm 362

## RESULT 179

AAP30469

ID AAP30469 standard; Protein: 7 AA.

XX AC AAP30469;

XX DT 31-MAY-1992 (first entry)

XX Sequence of polypeptide deriv. with antagonistic properties against  
DE substance P.

XX Substance P antagonist; chronic pain therapy; high blood pressure;  
KW hypertension; hypotensive agent.

XX Key Location/Qualifiers

FT Modified-site 1 /label= H-Pro

FT /note= "This AA may be omitted, in which case AA(2)  
FT - Prop. Gln"

FT Modified-site 3 /label= D-Trp

FT Modified-site 5 /label= D-Trp

FT Modified-site 7 /label= Met-NH2

XX DE3205991-A.

XX PN 01-SEP-1983.

XX PD 16-FEB-1983; 83DE-3467187.

XX PR 19-FEB-1982; 82DE-3205991.

XX PA (FERR-) FERRING ARZNEIMITTE.

XX PI Horig J, Schultheiss H;

XX DR WPI: 1983-753766/36.

XX Polypeptide derivs. contg. naturally occurring amino acids -  
PT antagonists against substance P and for treatment of pain and  
PT high blood pressure, including corneal inflammation  
XX Example; Page 16; 26pp; German.

XX The peptides of the invention can be used to treat chronic pain  
CC conditions and high blood pressure, including e.g. chronic  
CC inflammations of the cornea caused by various circumstances, e.g.  
CC lengthy exposure to UV light, IR-rays or chemicals.

SQ Sequence 7 AA;

Query Match 59.2%; Score 42; DB 4; Length 7;  
 Best Local Similarity 100.0%; Pred. NO. 4.3e+05;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 QWFWLM 11  
 |||||  
 Db 2 qwfwlm 7

RESULT 180

AAR21960  
 ID AAR21960 standard; Peptide; 11 AA.

AC AAR21960;

XX 25-JUN-1992 (first entry)

XX Cyclic substance P [Hcys 5,9].

XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
 KW syndrome; hereditary cerebral haemorrhage.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 5

FT /label= OTHER

FT /note= "OTHER = homocysteine"

FT Misc-difference 9

FT /label= OTHER

FT /note= "OTHER = homocysteine"

XX WO9202248-A.

XX 20-FEB-1992.

XX 29-JUL-1991; 9LWO-US05323.

XX 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using  
 PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
 PT B, for treating Alzheimer's disease, Downs syndrome, etc.

XX Claim 11; Page 22; 35pp; English.

XX The peptide is the tachykinin agonist, substance P with  
 CC homocysteine substituted at positions 5 and 9, with a disulphide  
 CC bond formed between them making the peptide cyclic. The  
 CC peptide was synthesised by standard solid phase synthesis.  
 CC Neuronal accumulation of beta-amyloid may be treated by administ-  
 CC ration of tachykinin agonists. The peptide can reduce the neuro-  
 CC toxic effects of a beta-amyloid related polypeptide on cultured  
 CC neurons. The peptide and its analogues are useful for controlling  
 CC diseases characterised by beta amyloid accumulation in the brain  
 CC such as Alzheimer's disease and Down's syndrome.  
 CC See also AAR21932-75.

XX Sequence 11 AA;

Query Match

Best Local Similarity 59.2%; Score 42; DB 13; Length 11;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
 |||||  
 Db 1 rpkipqffxlm 11

RESULT 181

AAR21939  
 ID AAR21939 standard; Protein; 11 AA.

AC AAR21939;

XX 25-JUN-1992 (first entry)

XX Substance P [Ile 8].

XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
 KW syndrome; hereditary cerebral haemorrhage.

XX Synthetic.

XX WO9202248-A.

XX 20-FEB-1992.

XX 29-JUL-1991; 9LWO-US05323.

XX 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using  
 PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
 PT B, for treating Alzheimer's disease, Downs syndrome, etc.

XX Claim 10; Page 21; 35pp; English.

XX The peptide is the tachykinin agonist substance P with an  
 CC isoleucine residue substituted at position 8. The peptide was  
 CC synthesised by standard solid phase synthesis. Neuronal  
 CC accumulation of beta-amyloid may be treated by administration of  
 CC tachykinin agonists. The peptide can reduce the neurotoxic effects  
 CC of a beta-amyloid related polypeptide on cultured neurons. The  
 CC peptide and its analogues are useful for controlling diseases  
 CC characterised by beta amyloid accumulation in the brain such as  
 CC Alzheimer's disease and Down's syndrome.  
 CC See also AAR21932-75.

XX Sequence 11 AA;

Query Match

Best Local Similarity 59.2%; Score 42; DB 13; Length 11;  
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
 |||||  
 Db 1 rpkipqffxlm 11

RESULT 182

AAR21949  
 ID AAR21949 standard; Protein; 11 AA.

AC AAR21949;

XX 25-JUN-1992 (first entry)

XX Substance P [Pro 3].



XX Disclosure; Column 17-18; 28pp; English.  
XX  
XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92673-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.  
XX Sequence 11 AA;  
SQ

Query Match 59.2%; Score 42; DB 20; Length 11;  
Best Local Similarity 72.7%; Pred. No. 2;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
DB 1 rpkpqffglm 11

RESULT 185  
AAW92673  
ID AAW92673 standard; peptide; 11 AA.  
XX  
AC AAW92673;  
XX  
DT 30-APR-1999 (first entry)  
XX  
DE Human tachykinin agonist beta-amyloid peptide fragment #19.  
XX  
XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX  
OS Homo sapiens.  
XX

PN US5876948-A.  
XX  
PD 02-MAR-1999.  
XX  
PF 27-JUL-1991; 91US-0737371.  
XX  
PR 29-JUL-1991; 91US-0737371.  
PR 27-JUL-1990; 90US-0559173.  
XX  
PA (CHIL-) CHILDRENS MEDICAL CENT.  
XX

PI Yankner BA;  
XX  
XX WPI; 1999-189630/16.  
XX  
XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX  
XX Disclosure; Column 17-18; 28pp; English.  
XX

XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.  
XX Sequence 11 AA;  
SQ

Query Match 59.2%; Score 42; DB 20; Length 11;  
Best Local Similarity 72.7%; Pred. No. 2;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
DB 1 rpppqffglm 11

RESULT 186  
AAB49755  
ID AAB49755 standard; peptide; 11 AA.  
XX  
AC AAB49755;  
XX  
DT 17-APR-2001 (first entry)  
XX  
DE Complex sugar bound peptide (SBP) amino acid sequence.  
XX  
KW Sugar peptide complex; SBP; sugar bound peptide; enzymatically stable.  
XX  
OS Synthetic.  
XX  
PN JP2000319297-A.  
XX  
PD 21-NOV-2000.  
XX  
PF 30-MAR-1999; 99JP-0088030.  
XX  
PR 30-MAR-1999; 99JP-0088030.  
XX  
PA (NOCK ) ZH NOGUCHI KENKYUSHO.  
XX  
DR WPI; 2001-184996/19.  
XX

XX A process for preparation of enzymically stable sugar peptide complex  
PT -  
XX  
XX Example 2; Page 3; 4pp; Japanese.  
XX  
XX This invention relates to a process for the preparation of an  
CC enzymatically stable sugar peptide complex, and includes an in vivo  
CC stable inhibitor of peptide-N-glycanase (EC. 3.5.1.52). The process can  
CC be used for the investigation of in vivo reciprocal recognition of  
CC cell-cell and substrate-receptor interaction, and their metabolism. The  
CC present sequence represents a complex sugar bound peptide (SBP) amino  
CC acid sequence prepared by the process of the invention.  
XX  
SQ Sequence 11 AA;

Query Match 59.2%; Score 42; DB 22; Length 11;  
Best Local Similarity 63.6%; Pred. No. 2;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKPQQWFWM 11  
DB 1 kprpqffglm 11

RESULT 187  
AAE04896  
ID AAE04896 standard; Protein; 398 AA.  
XX  
AC AAE04896;  
XX  
DT 10-SEP-2001 (first entry)  
XX  
DE Human transporter and ion channel-9 (TRICH-9) protein.  
XX  
KW Human; transporter and ion channel-9; TRICH-9; vaccine; cystic fibrosis;  
KW gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;

hypertension; angina; neurological disorder; asthma; bipolar disorder; dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia; Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke; Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy; demyelinating disease; mental disorder; Schizophrenia; polymyositis; muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease; dermatomyositis; diabetes mellitus; immunological disorder; psoriasis; rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus; sickle cell anaemia; Wilson's disease; infertility; Cushing's disease; scleroderma; pulmonary artery stenosis; neutropenic; Addison's disease; malabsorption syndrome; hypercholesterolaemia; cancer.

Homo sapiens.

Key . Location/Qualifiers  
Domain 217..242  
/label= Transmembrane\_domain  
Domain 247..264  
/label= Transmembrane\_domain  
Domain 350..368  
/label= Transmembrane\_domain

WO200146258-A2.

28-JUN-2001.

22-DEC-2000; 2000WO-US35095.

23-DEC-1999; 99US-0172000.

14-JAN-2000; 2000US-0176083.

21-JAN-2000; 2000US-0177332.

28-JAN-2000; 2000US-0178572.

02-FEB-2000; 2000US-0179758.

10-FEB-2000; 2000US-0181625.

(INCY-) INCYTE GENOMICS INC.

Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R, Lal P;

Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;

Tang YT, Khan FA;

WPI; 2001-418042/44.

N-PSDB; AAD09560.

Novel human transporter and ion channel proteins useful for treating

disorders

Claim 1; Page 121-122; 160pp; English.

The present sequence is transporter and ion channel-9 (TRICH-9) protein.

TRICH is used as vaccine. TRICH is useful for treating a disease or

condition associated with decreased expression of functional TRICH,

such as transport disorder including amyotrophic lateral sclerosis,

cystic fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth

disease, Duchenne muscular dystrophy, angina and hypertension,

neurological disorders including Alzheimer's disease, amnesia, bipolar

disorder, dementia depression, epilepsy, ischaemic cerebrovascular

disease, stroke, cerebral neoplasms, Pick's disease, Huntington's

disease and Parkinson's disease, demyelinating diseases, mental disorders

including mood, anxiety, Schizophrenia and seasonal affective disorder,

muscle disorder including cardiomyopathy, myocarditis, polymyositis,

dermatomyositis, arrhythmias and asthma and immunological disorders

including AIDS, adult respiratory distress syndrome (ARDS), allergies,

anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's

syndrome, systemic lupus erythematosus and other diseases including

sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary

artery stenosis, Grave's disease, Cushing's disease, Addison's disease,

glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers

psoriasis and viral, bacterial, fungal, helminthic and protozoal

infections. TRICH DNA is useful in gene therapy and in diagnostic

purposes.

SQ Sequence 398 AA;

Query Match 59.2%; Score 42; DB 22; Length 398;

Best Local Similarity 75.0%; Pred. No. 65;

Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 PQQWFWM 11

|||||;

Db 38 pqgyfwll 45

RESULT 188

AAW50976

ID AAW50976 standard; peptide; 8 AA.

XX AC AAW50976;

XX DT 31-JUL-1998 (first entry)

XX DE Substance P analogue residues 4-11, [D-Pro4,D-Trp7,9,10,Val8].

XX KW Vasoactive intestinal peptide; VIP; antagonist; somatostatin; bombesin;

KW KW Substance P; cancer; inhibition; growth hormone releasing factor;

KW KW spantide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Misc-difference 1 /note= "D-form residue"

FT Misc-difference 4 /note= "D-form residue"

FT Misc-difference 6 /note= "D-form residue"

FT Misc-difference 7 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Modified-site /note= "C-terminal amide"

XX PN EP835662-A2.

XX XX 15-APR-1998.

XX PF 11-DEC-1996; 96EP-0309012.

XX PR 08-OCT-1996; 96US-0727679.

XX PR 16-AUG-1996; 96IN-0001822.

XX PA (NAIM-) NAT INST IMMUNOLOGY.

XX PI Jaggi M, Mukherjee R;

XX DR WPI; 1998-208959/19.

XX PT Composition containing analogues of vasoactive intestinal peptide,

somatostatin - bombesin and substance P, for treatment of tumours

PT and for inhibiting over-expression of these peptide(s)

XX PS Disclosure; Page 13; 49pp; English.

XX CC The invention relates to a new composition which comprises: (i) the

somatostatin analogue SOM2 AGCKNFFQWKPTSDC (3-14 disulphide bridge),

and (ii) at least 4 of the peptides: antagonist of vasoactive

intestinal peptide (VIP1); VIP receptor-binding inhibitor (VIP2); VIP

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

CC inhibit or modulate over-expression of, e.g. VIP. A wide range of cancer  
CC cells express receptors for VIP, somatostatin, bombesin and/or substance  
CC P. The present sequence represents a substance P analogue.

XX  
SQ Sequence 8 AA;

Query Match 57.7%; Score 41; DB 19; Length 8;  
Best Local Similarity 75.0%; Pred. No. 4.3e+05;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PQQFWFLM 11  
| | | | |  
Db 1 pqqvwwvm 8

RESULT 189

AAW92711  
ID AAW92711 standard; peptide; 8 AA.

XX  
AC AAW92711;

XX  
DT 30-APR-1999 (first entry)

XX Human tachykinin agonist beta-amyloid peptide fragment #57.

XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congenital angiodysplasia.

XX Homo sapiens.

XX  
PN US5876948-A.

XX  
PD 02-MAR-1999.

XX  
PF 27-JUL-1991; 91US-0737371.

XX  
PR 29-JUL-1991; 91US-0737371.

XX  
PR 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MEDICAL CENT.

XX Yankner BA;

XX WPI; 1999-189630/16.

XX Screening for neurotoxin inhibitors - by testing compounds for their  
effect on beta-amyloid peptide neurotoxic effect on neuronal cells

XX Disclosure; Column 35-36; 28pp; English.

XX This invention describes a method for screening compounds for inhibiting  
a neurotoxin. The method involves incubating tachykinin agonists with  
neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
used for identifying compounds for treating diseases characterised by an  
undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
with amyloidosis and non-inherited congenital angiodysplasia with cerebral  
haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
beta-amyloid peptide fragments.

XX  
SQ Sequence 8 AA;

Query Match 57.7%; Score 41; DB 20; Length 8;  
Best Local Similarity 87.5%; Pred. No. 4.3e+05;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 8  
| | | | |  
Db 1 rpkpqgff 8

RESULT 190

AAR21932

ID AAR21932 standard; peptide; 9 AA.

XX  
AC AAR21932;

XX  
DT 25-JUN-1992 (first entry)

XX Substance P (1-9) fragment.

XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.

XX Synthetic.

XX WO9202248-A.

XX  
PD 20-FEB-1992.

XX  
PF 29-JUL-1991; 91WO-US05323.

XX  
PR 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using  
tachykinin agonists e.g. substance P, physalaemin and neurokinin  
B, for treating Alzheimer's disease, Downs syndrome, etc.

XX Claim 9; Page 21; 35pp; English.

XX The peptide is a tachykinin agonist consisting of residues 1-9 of  
substance P. The peptide was synthesised by standard solid phase  
synthesis. Analogues of the peptide, with C-terminal deletions down  
to substance P (1-4) were also synthesised. Neuronal accumulation of  
beta-amyloid may be treated by administration of these tachykinin  
agonists. The peptides reduce the neurotoxic effects of a beta-  
amyloid related polypeptide on cultured neurons. The peptide and  
its analogues are useful for controlling diseases characterised by  
beta amyloid accumulation in the brain such as Alzheimer's disease  
and Down's syndrome.

XX See also AAR21933-75.

XX  
SQ Sequence 9 AA;

Query Match 57.7%; Score 41; DB 13; Length 9;  
Best Local Similarity 87.5%; Pred. No. 4.3e+05;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 8  
| | | | |  
Db 1 rpkpqgff 8

RESULT 191

AAV03162

ID AAV03162 standard; peptide; 9 AA.

XX  
AC AAV03162;

XX  
DT 10-JUN-1999 (first entry)

XX Substance P fragment P/1-9#.

XX Opioid peptide; opioid analgesia; enhancer; opioid anaesthesia;  
KW substance P.

[illegible]



CC The targeting moiety comprises a light chain and an amine end segment of  
CC a heavy chain and comprises Substance P as the targeting moiety. The pain  
CC alleviating effects persist for 2-6 months. The present sequence  
CC represents a substance P fragment, and is used in the course of the  
CC invention.  
XX  
SQ Sequence 9 AA;

Query Match 57.7%; Score 41; DB 22; Length 9;  
Best Local Similarity 87.5%; Pred. No. 4.3e+05;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPKPOQWF 8  
| | | | | | | |  
Db 1 rpkpqgff 8

RESULT 194  
AAB98878  
ID AAB98878 standard; Peptide; 9 AA.

XX AC AAB98878;

XX 14-AUG-2001 (first entry)

XX DE Chimeric analgesic peptide #34.

XX Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.  
XX  
OS Synthetic.

XX FH Key Location/Qualifiers  
FT Modified-site 9  
FT /label= OTHER  
FT /note= "C-terminal amide"

XX WO200130371-A2.

XX PD 03-MAY-2001.

XX 27-OCT-2000; 2000WO-US29789.

XX 28-OCT-1999; 99US-0428692.

XX (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;

XX WPI; 2001-397593/42.

XX New chimeric peptides used for treating pain comprise opioid receptor  
PT binding group and nociceptive receptor binding group -  
PS

XX Claim 10; Page 15; 34pp; English.

XX The present invention describes a number of chimeric peptides comprising  
CC an opioid receptor binding moiety and a nociceptive receptor binding  
CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
CC opioid receptor based peptides alone, tolerance does not result from  
CC their long-term use. The present sequence is one of the peptides of the  
CC invention.  
XX

SQ Sequence 9 AA;

Query Match 57.7%; Score 41; DB 22; Length 9;  
Best Local Similarity 87.5%; Pred. No. 4.3e+05;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPKPOQWF 8  
| | | | | | | |

Db 1 rpkpqgff 8  
RESULT 195  
AAB91444  
ID AAB91444 standard; Peptide; 9 AA.

XX AC AAB91444;

XX 22-JUN-2001 (first entry)

XX Tachykinins peptide SEQ ID NO:620.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX

OS Homo sapiens.  
OS Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US13576.

XX 17-MAY-1999; 99US-0134406.

XX 10-SEP-1999; 99US-0134406.

XX 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT

XX Disclosure; Page 401; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptide stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX

SQ Sequence 9 AA;

Query Match 57.7%; Score 41; DB 22; Length 9;  
Best Local Similarity 87.5%; Pred. No. 4.3e+05;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPKPOQWF 8  
| | | | | | | |

Db 1 rpkpqgff 8

RESULT 196

AAB91410  
ID AAB91410 standard; Peptide; 10 AA.  
XX  
XX AC AAB91410;  
XX  
XX DT 22-JUN-2001 (first entry)  
XX  
XX DE Tachykinins peptide SEQ ID NO:586.  
XX  
XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
XX OS Homo sapiens.  
OS Synthetic.  
XX  
XX PN WO200069900-A2.  
XX  
XX PD 23-NOV-2000.  
XX  
XX PF 17-MAY-2000; 2000WO-US13576.  
XX  
XX PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
XX PA (CONJ-) CONJUCHEM INC.  
XX  
XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX  
XX DR WPI; 2001-112059/12.  
XX  
XX PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
XX PS Disclosure; Page 391; 733pp; English.  
XX  
XX CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
XX SQ Sequence 10 AA;  
  
Query Match 57.7%; Score 41; DB 22; Length 10;  
Best Local Similarity 87.5%; Pred. No. 2.6;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RPKPQOWF 8  
Db 1 rpkpqgff 8  
  
RESULT 197  
AAB91422  
ID AAB91422 standard; Peptide; 10 AA.  
XX  
XX AC AAB91422;  
XX  
XX DT 22-JUN-2001 (first entry)  
XX  
XX DE Tachykinins peptide SEQ ID NO:608.

XX  
XX DT 22-JUN-2001 (first entry)  
XX  
XX DE Tachykinins peptide SEQ ID NO:598.  
XX  
XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
XX OS Homo sapiens.  
OS Synthetic.  
XX  
XX PN WO200069900-A2.  
XX  
XX PD 23-NOV-2000.  
XX  
XX PF 17-MAY-2000; 2000WO-US13576.  
XX  
XX PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
XX PA (CONJ-) CONJUCHEM INC.  
XX  
XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX  
XX DR WPI; 2001-112059/12.  
XX  
XX PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
XX PS Disclosure; Page 395; 733pp; English.  
XX  
XX CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
XX SQ Sequence 10 AA;  
  
Query Match 57.7%; Score 41; DB 22; Length 10;  
Best Local Similarity 87.5%; Pred. No. 2.6;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RPKPQOWF 8  
Db 1 rpkpqgff 8  
  
RESULT 198  
AAB91432  
ID AAB91432 standard; Peptide; 10 AA.  
XX  
XX AC AAB91432;  
XX  
XX DT 22-JUN-2001 (first entry)  
XX  
XX DE Tachykinins peptide SEQ ID NO:608.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
 KW blood component; modification; succinimidyl; maleimido group; amino;  
 KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
 XX Synthetic.  
 OS Homo sapiens.  
 OS WO200069900-A2.  
 PN 23-NOV-2000.  
 PD 17-MAY-2000; 2000WO-US13576.  
 PF 17-MAY-1999; 99US-0134406.  
 XX 10-SEP-1999; 99US-0153406.  
 PR 15-OCT-1999; 99US-0159783.  
 XX (CONJ-) CONJUCHEM INC.  
 PA Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
 XX WPI; 2001-112059/12.  
 DR Modifying and attaching therapeutic peptides to albumin prevents  
 XX peptidase degradation, useful for increasing length of in vivo activity  
 PT  
 PT Disclosure; Page 398; 733pp; English.  
 XX The present invention describes a modified therapeutic peptide (I)  
 CC comprising a therapeutically active amino acid region (III) and a  
 CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
 CC a less therapeutically active amino acid region (IV), which covalently  
 CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
 CC peptide stabilised therapeutic peptide composed of 3-50 amino acids.  
 CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
 CC factors and neurotransmitters, to protect them from peptidase activity  
 CC in vivo for the treatment of various disorders. Endogenous therapeutic  
 CC peptides are not suitable as drug candidates as they require frequent  
 CC administration due to rapid degradation by peptidases in the body.  
 CC Modifying and attaching therapeutic peptides to albumin prevents or  
 CC reduces the action of peptidases to increase length of activity (half  
 CC life) and specificity as bonding to large molecules decreases  
 CC intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.  
 XX SQ Sequence 10 AA;

Query Match 57.7%; Score 41; DB 22; Length 10;  
 Best Local Similarity 87.5%; Pred. No. 2.6;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 8  
 (|||||:|  
 Db 1 rpkpqgff 8

RESULT 199  
 AAR21940  
 ID AAR21940 standard; Protein; 11 AA.  
 XX  
 AC AAR21940;  
 XX  
 DT 25-JUN-1992 (first entry)  
 XX  
 DE Substance P (Pro 10).  
 XX  
 KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
 KW syndrome; hereditary cerebral haemorrhage.

OS Synthetic.  
 XX WO9202248-A.  
 PN 20-FEB-1992.  
 PD 29-JUL-1991; 91WO-US05323.  
 XX 27-JUL-1990; 90US-0559173.  
 PR (CHIL-) CHILDRENS MED CENT.  
 PA Yankner BA;  
 XX WPI; 1992-079804/10.  
 DR Treatment of neuronal accumulation of beta-amyloid - using  
 XX tachykinin agonists e.g. substance P, physalaemin and neurokinin  
 PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
 XX Claim 10; Page 21; 35pp; English.  
 PS The peptide is the tachykinin agonist substance P with a Proline  
 XX residue substituted at position 10. The peptide was  
 CC synthesised by standard solid phase synthesis. Neuronal  
 CC accumulation of beta-amyloid may be treated by administration of  
 CC tachykinin agonists. The peptide can reduce the neurotoxic effects  
 CC of a beta-amyloid related polypeptide on cultured neurons. The  
 CC peptide and its analogues are useful for controlling diseases  
 CC characterised by beta amyloid accumulation in the brain such as  
 CC Alzheimer's disease and Down's syndrome.  
 CC See also AAR21932-75.  
 XX SQ Sequence 11 AA;

Query Match 57.7%; Score 41; DB 13; Length 11;  
 Best Local Similarity 87.5%; Pred. No. 2.9;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 8  
 (|||||:|  
 Db 1 rpkpqgff 8

RESULT 200  
 AAW92716  
 ID AAW92716 standard; peptide; 11 AA.  
 XX  
 AC AAW92716;  
 XX  
 DT 30-APR-1999 (first entry)  
 XX  
 DE Human tachykinin agonist beta-amyloid peptide fragment #62.  
 XX  
 KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
 KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
 KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
 XX  
 OS Homo sapiens.  
 XX US5876948-A.  
 PN 02-MAR-1999.  
 PD 27-JUL-1991; 91US-0737371.  
 XX 29-JUL-1991; 91US-0737371.  
 PR 27-JUL-1990; 90US-0559173.  
 XX (CHIL-) CHILDRENS MEDICAL CENT.  
 PA Yankner BA;  
 PI

XX WPI; 1999-189630/16.  
 XX Screening for neurotoxin inhibitors - by testing compounds for their  
 PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
 XX  
 XX PS Disclosure; Column 37-38; 28pp; English.  
 XX  
 CC This invention describes a method for screening compounds for inhibiting  
 CC a neurotoxin. The method involves incubating tachykinin agonists with  
 CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
 CC used for identifying compounds for treating diseases characterised by an  
 CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
 CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
 CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
 CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
 CC beta-amyloid peptide fragments.  
 XX  
 XX Sequence 11 AA;

Query Match 57.7%; Score 41; DB 20; Length 11;  
 Best Local Similarity 87.5%; Pred. No. 2.9;  
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RPKPQWF 8  
 |||||  
 Db 1 rpkpqff 8

Search completed: April 1, 2002, 16:18:23  
 Job time: 54 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2002, 16:18:44 ; Search time 20.12 Seconds  
(without alignments)

12.303 Million cell updates/sec

Title: US-09-988-792-2

Perfect score: 71

Sequence: 1 RPKPQOWFLM 11

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 192

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 50%

Maximum Match 100%

Listing first 1000 summaries

Database :

Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*

2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*

3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*

4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*

5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*

6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description       |
|------------|-------|-------------|--------|-------|-------------------|
| 1          | 71    | 100.0       | 11     | 6     | 5441935-5         |
| 2          | 68    | 95.8        | 11     | 6     | 5441935-2         |
| 3          | 61    | 85.9        | 11     | 6     | 5441935-3         |
| 4          | 54    | 76.1        | 8      | 6     | 5441935-10        |
| 5          | 49    | 69.0        | 11     | 2     | US-07-737-371E-12 |
| 6          | 48    | 67.6        | 8      | 6     | 5441935-6         |
| 7          | 48    | 67.6        | 11     | 1     | US-07-934-553-1   |
| 8          | 48    | 67.6        | 11     | 1     | US-08-184-935-12  |
| 9          | 48    | 67.6        | 11     | 1     | US-08-269-288-2   |
| 10         | 48    | 67.6        | 11     | 1     | US-08-338-484-1   |
| 11         | 48    | 67.6        | 11     | 1     | US-08-175-432-1   |
| 12         | 48    | 67.6        | 11     | 1     | US-08-225-474-1   |
| 13         | 48    | 67.6        | 11     | 1     | US-08-391-910-2   |
| 14         | 48    | 67.6        | 11     | 1     | US-08-418-994-2   |
| 15         | 48    | 67.6        | 11     | 1     | US-08-480-505-3   |
| 16         | 48    | 67.6        | 11     | 1     | US-08-391-814-2   |
| 17         | 48    | 67.6        | 11     | 1     | US-08-167-870-1   |
| 18         | 48    | 67.6        | 11     | 1     | US-08-255-272-6   |
| 19         | 48    | 67.6        | 11     | 1     | US-08-441-591-6   |
| 20         | 48    | 67.6        | 11     | 1     | US-08-303-362A-6  |
| 21         | 48    | 67.6        | 11     | 1     | US-08-462-859A-1  |
| 22         | 48    | 67.6        | 11     | 1     | US-08-123-659A-1  |
| 23         | 48    | 67.6        | 11     | 1     | US-08-462-415-2   |
| 24         | 48    | 67.6        | 11     | 1     | US-08-463-874-2   |
| 25         | 48    | 67.6        | 11     | 1     | US-08-464-247A-1  |
| 26         | 48    | 67.6        | 11     | 1     | US-08-464-248A-1  |
| 27         | 48    | 67.6        | 11     | 1     | US-08-444-135-2   |

|    |                    |                    |
|----|--------------------|--------------------|
| 11 | US-08-318-391-2    | Sequence 2, Appli  |
| 11 | US-08-796-598-11   | Sequence 11, Appl  |
| 11 | US-08-447-175A-11  | Sequence 11, Appl  |
| 11 | US-07-737-371E-21  | Sequence 21, Appl  |
| 11 | US-07-737-371E-22  | Sequence 22, Appl  |
| 11 | US-07-737-371E-23  | Sequence 23, Appl  |
| 11 | US-07-737-371E-24  | Sequence 24, Appl  |
| 11 | US-07-737-371E-25  | Sequence 25, Appl  |
| 11 | US-07-737-371E-26  | Sequence 26, Appl  |
| 11 | US-07-737-371E-27  | Sequence 27, Appl  |
| 11 | US-07-737-371E-28  | Sequence 28, Appl  |
| 11 | US-07-737-371E-29  | Sequence 29, Appl  |
| 11 | US-07-737-371E-30  | Sequence 30, Appl  |
| 11 | US-07-737-371E-31  | Sequence 31, Appl  |
| 11 | US-07-737-371E-32  | Sequence 32, Appl  |
| 11 | US-07-737-371E-33  | Sequence 33, Appl  |
| 11 | US-07-737-371E-34  | Sequence 34, Appl  |
| 11 | US-07-737-371E-35  | Sequence 35, Appl  |
| 11 | US-07-737-371E-36  | Sequence 36, Appl  |
| 11 | US-07-737-371E-37  | Sequence 37, Appl  |
| 11 | US-07-737-371E-38  | Sequence 38, Appl  |
| 11 | US-07-737-371E-39  | Sequence 39, Appl  |
| 11 | US-07-737-371E-40  | Sequence 40, Appl  |
| 11 | US-07-737-371E-41  | Sequence 41, Appl  |
| 11 | US-07-737-371E-42  | Sequence 42, Appl  |
| 11 | US-07-737-371E-43  | Sequence 43, Appl  |
| 11 | US-07-737-371E-44  | Sequence 44, Appl  |
| 11 | US-07-737-371E-45  | Sequence 45, Appl  |
| 11 | US-07-737-371E-46  | Sequence 46, Appl  |
| 11 | US-07-737-371E-47  | Sequence 47, Appl  |
| 11 | US-07-737-371E-48  | Sequence 48, Appl  |
| 11 | US-07-737-371E-49  | Sequence 49, Appl  |
| 11 | US-07-737-371E-50  | Sequence 50, Appl  |
| 11 | US-07-737-371E-51  | Sequence 51, Appl  |
| 11 | US-07-737-371E-52  | Sequence 52, Appl  |
| 11 | US-07-737-371E-53  | Sequence 53, Appl  |
| 11 | US-07-737-371E-54  | Sequence 54, Appl  |
| 11 | US-07-737-371E-55  | Sequence 55, Appl  |
| 11 | US-07-737-371E-56  | Sequence 56, Appl  |
| 11 | US-07-737-371E-57  | Sequence 57, Appl  |
| 11 | US-07-737-371E-58  | Sequence 58, Appl  |
| 11 | US-07-737-371E-59  | Sequence 59, Appl  |
| 11 | US-07-737-371E-60  | Sequence 60, Appl  |
| 11 | US-07-737-371E-61  | Sequence 61, Appl  |
| 11 | US-07-737-371E-62  | Sequence 62, Appl  |
| 11 | US-07-737-371E-63  | Sequence 63, Appl  |
| 11 | US-07-737-371E-64  | Sequence 64, Appl  |
| 11 | US-07-737-371E-65  | Sequence 65, Appl  |
| 11 | US-07-737-371E-66  | Sequence 66, Appl  |
| 11 | US-07-737-371E-67  | Sequence 67, Appl  |
| 11 | US-07-737-371E-68  | Sequence 68, Appl  |
| 11 | US-07-737-371E-69  | Sequence 69, Appl  |
| 11 | US-07-737-371E-70  | Sequence 70, Appl  |
| 11 | US-07-737-371E-71  | Sequence 71, Appl  |
| 11 | US-07-737-371E-72  | Sequence 72, Appl  |
| 11 | US-07-737-371E-73  | Sequence 73, Appl  |
| 11 | US-07-737-371E-74  | Sequence 74, Appl  |
| 11 | US-07-737-371E-75  | Sequence 75, Appl  |
| 11 | US-07-737-371E-76  | Sequence 76, Appl  |
| 11 | US-07-737-371E-77  | Sequence 77, Appl  |
| 11 | US-07-737-371E-78  | Sequence 78, Appl  |
| 11 | US-07-737-371E-79  | Sequence 79, Appl  |
| 11 | US-07-737-371E-80  | Sequence 80, Appl  |
| 11 | US-07-737-371E-81  | Sequence 81, Appl  |
| 11 | US-07-737-371E-82  | Sequence 82, Appl  |
| 11 | US-07-737-371E-83  | Sequence 83, Appl  |
| 11 | US-07-737-371E-84  | Sequence 84, Appl  |
| 11 | US-07-737-371E-85  | Sequence 85, Appl  |
| 11 | US-07-737-371E-86  | Sequence 86, Appl  |
| 11 | US-07-737-371E-87  | Sequence 87, Appl  |
| 11 | US-07-737-371E-88  | Sequence 88, Appl  |
| 11 | US-07-737-371E-89  | Sequence 89, Appl  |
| 11 | US-07-737-371E-90  | Sequence 90, Appl  |
| 11 | US-07-737-371E-91  | Sequence 91, Appl  |
| 11 | US-07-737-371E-92  | Sequence 92, Appl  |
| 11 | US-07-737-371E-93  | Sequence 93, Appl  |
| 11 | US-07-737-371E-94  | Sequence 94, Appl  |
| 11 | US-07-737-371E-95  | Sequence 95, Appl  |
| 11 | US-07-737-371E-96  | Sequence 96, Appl  |
| 11 | US-07-737-371E-97  | Sequence 97, Appl  |
| 11 | US-07-737-371E-98  | Sequence 98, Appl  |
| 11 | US-07-737-371E-99  | Sequence 99, Appl  |
| 11 | US-07-737-371E-100 | Sequence 100, Appl |



Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFWLM 11  
|||||  
Db 1 RPKPQQFWLL 11

RESULT 3  
5441935-3  
; Patent No. 5441935  
; APPLICANT: Rozenqurt, Enrique; Zachary, Ian; Woll, Penella  
; TITLE OF INVENTION: ROTH FACTOR RECEPTORS  
; NUMBER OF SEQUENCES:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/939,587  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 814,064  
; FILING DATE: 23-DEC-1991  
; APPLICATION NUMBER: 411,536  
; FILING DATE: 29-NOV-1989  
; SEQ ID NO:3:  
; LENGTH: 11  
5441935-3

Query Match 85.9%; Score 61; DB 6; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.0012;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFWLM 11  
|||||  
Db 1 RPKPQQFWLM 11

RESULT 4  
5441935-10  
; Patent No. 5441935  
; APPLICANT: Rozenqurt, Enrique; Zachary, Ian; Woll, Penella  
; TITLE OF INVENTION: ROTH FACTOR RECEPTORS  
; NUMBER OF SEQUENCES:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/939,587  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 814,064  
; FILING DATE: 23-DEC-1991  
; APPLICATION NUMBER: 411,536  
; FILING DATE: 29-NOV-1989  
; SEQ ID NO:10:  
; LENGTH: 8  
5441935-10

Query Match 76.1%; Score 54; DB 6; Length 8;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 PQQFWLM 11  
|||||  
Db 1 PQQFWLM 8

RESULT 5  
US-07-737-371E-12  
; Sequence 12, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-737-371E-12

Query Match 69.0%; Score 49; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.072;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFWLM 11  
|||||  
Db 1 RPKPQYFGLM 11

RESULT 6  
5441935-6  
; Patent No. 5441935  
; APPLICANT: Rozenqurt, Enrique; Zachary, Ian; Woll, Penella  
; TITLE OF INVENTION: ROTH FACTOR RECEPTORS  
; NUMBER OF SEQUENCES:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/939,587  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 814,064  
; FILING DATE: 23-DEC-1991  
; APPLICATION NUMBER: 411,536  
; FILING DATE: 29-NOV-1989  
; SEQ ID NO:6:  
; LENGTH: 8  
5441935-6

Query Match 67.6%; Score 48; DB 6; Length 8;  
Best Local Similarity 87.5%; Pred. No. 1.6e+05;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQFWLM 11  
|||||  
Db 1 PQQFWLM 8

RESULT 7

US-07-934-553-1  
; Sequence 1, Application US/07934553  
; Patent No. 5314690  
; GENERAL INFORMATION:  
; APPLICANT: PATTERSON, ROY  
; APPLICANT: HARRIS, KATHLEEN E  
; TITLE OF INVENTION: METHOD AND COMPOSITION FOR REDUCING Ige  
; TITLE OF INVENTION: ANTIBODIES TO SPECIFIC ALLERGENS  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TILTON, FALLON, LUNGWUS & CHESTNUT  
; STREET: 100 SOUTH WACKER DRIVE  
; CITY: CHICAGO  
; STATE: ILLINOIS  
; COUNTRY: USA  
; ZIP: 60606-4002  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/07/934,553  
; APPLICATION DATE: 19920821  
; FILING DATE: 19920821  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/705,071  
; FILING DATE: 24-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FENTRESS, SUSAN B  
; REGISTRATION NUMBER: 31,327  
; REFERENCE/DOCKET NUMBER: NU-90333CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/456-8000  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: unknown  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-07-934-553-1

Query Match 67.68; Score 48; DB 1; Length 11;  
Best Local Similarity 81.88; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
| | | | | | | | | | | |  
Db 1 RPKPQQFFGLM 11

RESULT 8  
US-08-184-935-12  
; Sequence 12, Application US/08184935  
; Patent No. 5476770  
; GENERAL INFORMATION:  
; APPLICANT: PRADELLES, PHILIPPE  
; TITLE OF INVENTION: IMMUNOMETRIC DETERMINATION OF AN ANTIGEN  
; TITLE OF INVENTION: OR HAPTEN  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/184,935  
; FILING DATE: 24-JAN-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NO. 5476770man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 846-286-0  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 11  
; OTHER INFORMATION: /note= "C-terminal amide"  
; US-08-184-935-12

Query Match 67.68; Score 48; DB 1; Length 11;  
Best Local Similarity 81.88; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
| | | | | | | | | | | |  
Db 1 RPKPQQFFGLM 11

RESULT 9  
US-08-269-288-2  
; Sequence 2, Application US/08269288  
; Patent No. 5491140  
; GENERAL INFORMATION:  
; APPLICANT: Bruns, Robert F.  
; APPLICANT: Gehlert, Donald R.  
; APPLICANT: Howbert, James J.  
; APPLICANT: Lunn, William H.W.  
; TITLE OF INVENTION: NAPHTHYL TACHYKININ RECEPTOR ANTAGONISTS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center/1104  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: United States of America  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/269,288  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gaylo, Paul J.  
; REGISTRATION NUMBER: 36,808  
; REFERENCE/DOCKET NUMBER: X-9715  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317) 276-0756  
; TELEFAX: (317) 276-3861  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids



; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-269-288-2

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWFWM 11  
|||||:|  
Db 1 RPKPOQFFGLM 11

## RESULT 10

US-08-338-484-1  
; Sequence 1, Application US/08338484  
; Patent No. 5494926

## ; GENERAL INFORMATION:

; APPLICANT: Owens, Andrew P.  
; APPLICANT: Teall, Martin R.  
; APPLICANT: Williams, Brian J.  
; TITLE OF INVENTION: 2/3-(HETEROCYCLIC ALKYL  
; TITLE OF INVENTION: AMINO)-1-(SUBSTITUTED PHENYL-METHOXY)-ETHANES/PROPANES AS  
; TITLE OF INVENTION: TACHYKININ RECEPTOR ANTAGONISTS  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dr. Robert J. No. 5494926th  
; STREET: 126 E. Lincoln Ave., P.O. Box 2000  
; CITY: Rahway  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07065-0900

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/338,484  
; FILING DATE: 18-NOV-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5494926th, Robert J.  
; REGISTRATION NUMBER: 27,366  
; REFERENCE/DOCKET NUMBER: T-1158  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 594-7262  
; TELEFAX: (908) 594-4720  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-338-484-1

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWFWM 11  
|||||:|  
Db 1 RPKPOQFFGLM 11

## RESULT 11

US-08-175-432-1  
; Sequence 1, Application US/08175432

; Patent No. 5495047  
; GENERAL INFORMATION:  
; APPLICANT: Saari, Walfred S. B.  
; APPLICANT: Van Niel, Monique B.  
; APPLICANT: Williams, Brian J.  
; TITLE OF INVENTION: FUSED TRICYCLIC COMPOUNDS,  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR USE  
; TITLE OF INVENTION: IN THERAPY  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NORTH, ROBERT J.  
; STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
; CITY: Rahway  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/175,432  
; FILING DATE: 07-JAN-1994  
; CLASSIFICATION: 560  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5495047th, Robert J.  
; REGISTRATION NUMBER: 27,366  
; REFERENCE/DOCKET NUMBER: T-1152Y  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 594-7262  
; TELEFAX: (908) 594-4720  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-175-432-1

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWFWM 11  
|||||:|  
Db 1 RPKPOQFFGLM 11

## RESULT 12

US-08-225-474-1  
; Sequence 1, Application US/08225474  
; Patent No. 5360915

## ; GENERAL INFORMATION:

; APPLICANT: Patterson, Roy  
; APPLICANT: Harris, Kathleen E.  
; TITLE OF INVENTION: Method and Composition for Treating  
; TITLE OF INVENTION: Ige Mediated Allergies  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Tilton, Fallon, Lungmus & Chestnut  
; STREET: 100 S. Wacker Drive, Suite 960  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-4002

## COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/225,474  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/934,553  
FILING DATE: 21-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/705,071  
FILING DATE: 24-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Tilton, Timothy L.  
REGISTRATION NUMBER: 16,926  
REFERENCE/DOCKET NUMBER: NU 9033-CIP2  
TELEPHONE: (312)-456-8000  
TELEFAX: (312)-456-7776  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-225-474-1

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQQFFGLM 11

RESULT 13  
US-08-391-910-2  
Sequence 2, Application US/08391910  
Patent No. 5563133  
GENERAL INFORMATION:  
APPLICANT: Hipskind, Philip A.  
TITLE OF INVENTION: HEXAMETHYLENIMINYL TACHYKININ RECEPTOR  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: United States of America  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Gaylo, Paul J.  
REGISTRATION NUMBER: 36,808  
REFERENCE/DOCKET NUMBER: X-9979  
TELEPHONE: (317) 276-0756  
TELEFAX: (317) 276-3861  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-391-910-2

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQQFFGLM 11

RESULT 14  
US-08-418-994-2  
Sequence 2, Application US/08418994  
Patent No. 5565568  
GENERAL INFORMATION:  
APPLICANT: Cho, Sung-Yong S.  
APPLICANT: Hipskind, Philip A.  
APPLICANT: Howbert, J. J.  
APPLICANT: Muehl, Brian S.  
APPLICANT: Nixon, James A.  
TITLE OF INVENTION: 2-ACYLAMINOPROPANAMIDES AS TACHYKININ  
TITLE OF INVENTION: RECEPTOR ANTAGONISTS  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Eli Lilly and Company  
STREET: Lilly Corporate Center  
CITY: Indianapolis  
STATE: Indiana  
COUNTRY: United States of America  
ZIP: 46285  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Gaylo, Paul J.  
REGISTRATION NUMBER: 36,808  
REFERENCE/DOCKET NUMBER: X-8252  
TELEPHONE: (317) 276-0756  
TELEFAX: (317) 276-3861  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-418-994-2

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQQFFGLM 11

RESULT 15  
US-08-480-505-3

; Sequence 3, Application US/08480505  
; Patent No. 5601821  
; GENERAL INFORMATION:  
; APPLICANT: STANNORTH, DENIS R  
; APPLICANT: LEWIN, IAN V  
; APPLICANT: NAYYAR, SARITA  
; APPLICANT: JONES, VALERIE  
; TITLE OF INVENTION: IMMUNOACTIVE PEPTIDES AND ANTIBODIES AND  
; TITLE OF INVENTION: THEIR USE IN ANTI-ALLERGY TREATMENT  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NIXON & VANDERHYE P.C.  
; STREET: 14TH FLOOR, 2200 CLARENDON BOULEVARD  
; CITY: ARLINGTON  
; STATE: VIRGINIA  
; COUNTRY: USA  
; ZIP: 22201-3360  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/480,505  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION NUMBER: US/08/102,692  
; FILING DATE:  
; APPLICATION NUMBER: US 07/776,380  
; FILING DATE: 26-NOV-1991  
; APPLICATION NUMBER: GB 8913737.6  
; FILING DATE: 15-JUN-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/GB90/00926  
; FILING DATE: 15-JUN-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MITCHARD, LEONARD C  
; REGISTRATION NUMBER: 29,009  
; REFERENCE/DOCKET NUMBER: 604-176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 875-0400  
; TELEFAX: (703) 525-3468  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; FRAGMENT TYPE: C-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Neuropeptide "Substance P"  
; US-08-480-505-3

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
|||||:| |  
DB 1 RPKPOQFFGLM 11

RESULT 16  
US-08-391-814-2  
; Sequence 2, Application US/08391814  
; Patent No. 5607947  
; GENERAL INFORMATION:  
; APPLICANT: Hipskind, Phillip A.  
; TITLE OF INVENTION: PYRROLIDINYL TACHYKININ RECEPTOR  
; TITLE OF INVENTION: ANTAGONISTS

; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: United States of America  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/391,814  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gaylo, Paul J.  
; REGISTRATION NUMBER: 36,808  
; REFERENCE/DOCKET NUMBER: X-9965  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317) 276-0736  
; TELEFAX: (317) 276-3861  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-391-814-2

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
|||||:| |  
DB 1 RPKPOQFFGLM 11

RESULT 17  
US-08-167-870-1  
; Sequence 1, Application US/08167870  
; Patent No. 5610183  
; GENERAL INFORMATION:  
; APPLICANT: OWENS, ANDREW P.  
; APPLICANT: WILLIAMS, BRIAN J.  
; TITLE OF INVENTION: AROMATIC COMPOUNDS, COMPOSITIONS  
; TITLE OF INVENTION: CONTAINING THEM AND THEIR USE IN THERAPY  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ROBERT J. NORTH  
; CITY: RAHWAY  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 07065  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/167,870  
; FILING DATE: 17-DEC-1993  
; CLASSIFICATION: 544  
; ATTORNEY/AGENT INFORMATION:  
; NAME: NORTH, ROBERT J.  
; REGISTRATION NUMBER: 27,366  
; REFERENCE/DOCKET NUMBER: T-1151Y

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908)594-7262  
TELEFAX: (908)594-4720  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-167-870-1

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 1 RPKPQQFFGLM 11

RESULT 18  
US-08-255-272-6  
Sequence 6, Application US/08255272  
Patent No. 5627268  
GENERAL INFORMATION:  
APPLICANT: Kumar, Ramesh  
APPLICANT: Sharma, Ajay  
APPLICANT: Khoury-Christianson, Anastasia  
APPLICANT: M.  
TITLE OF INVENTION: Production of Therapeutic Peptides in  
TITLE OF INVENTION: Transgenic Animals as a Fusion with Hemoglobin  
NUMBER OF SEQUENCES: 32  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/255,272  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
CLASSIFICATION: 435  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30742  
REFERENCE/DOCKET NUMBER: 6794-032  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-255-272-6

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 1 RPKPQQFFGLM 11

RESULT 19  
US-08-441-591-6  
Sequence 6, Application US/08441591  
Patent No. 5637682  
GENERAL INFORMATION:  
APPLICANT: NIEUWLANDT, D., GOLD, L. AND WECKER, M.  
TITLE OF INVENTION: HIGH-AFFINITY  
TITLE OF INVENTION: OLIGONUCLEOTIDE LIGANDS  
TITLE OF INVENTION: TO THE TACHYKININ  
TITLE OF INVENTION: SUBSTANCE P  
NUMBER OF SEQUENCES: 66  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Swanson & Bratschun, L.L.C.  
STREET: 8400 E. Prentice Avenue, Suite 200  
CITY: Englewood  
STATE: Colorado  
COUNTRY: USA  
ZIP: 80111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage  
COMPUTER: IBM compatible  
OPERATING SYSTEM: MS-DOS  
SOFTWARE: WordPerfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/441,591  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/303,362  
FILING DATE: 9-SEPTEMBER-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/714,131  
FILING DATE: 10-JUNE-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/931,473  
FILING DATE: 17-AUGUST-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/117,991  
FILING DATE: 8-SEPTEMBER 1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/536,428  
FILING DATE: 11-JUNE-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/964,624  
FILING DATE: 21-OCTOBER-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Barry J. Swanson  
REGISTRATION NUMBER: 33,215  
REFERENCE/DOCKET NUMBER: NEX21/C  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 793-3333  
TELEFAX: (303) 793-3433  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-441-591-6

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |

Db 1 RPKPQQFFGLM 11

## RESULT 20

US-08-303-362A-6

; Sequence 6, Application US/08303362A

; Patent No. 5648214

; GENERAL INFORMATION:

; APPLICANT: NIEUWLANDT, D., GOLD, L. AND WECKER, M.

; TITLE OF INVENTION: HIGH-AFFINITY

; TITLE OF INVENTION: OLIGONUCLEOTIDE LIGANDS

; TITLE OF INVENTION: TO THE TACHYKININ

; TITLE OF INVENTION: SUBSTANCE P

; NUMBER OF SEQUENCES: 66

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Swanson & Bratschun, L.L.C.

; STREET: 8400 E. Prentice Avenue, Suite 200

; CITY: Englewood

; STATE: Colorado

; COUNTRY: USA

; ZIP: 80111

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage

; COMPUTER: IBM compatible

; OPERATING SYSTEM: MS-DOS

; SOFTWARE: WordPerfect 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/303,362A

; FILING DATE: 9-SEPTEMBER-1994

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/714,131

; FILING DATE: 10-JUNE-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/931,473

; FILING DATE: 17-AUGUST-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/117,991

; FILING DATE: 8-SEPTEMBER 1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/536,428

; FILING DATE: 11-JUNE-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/964,624

; FILING DATE: 21-OCTOBER-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Barry J. Swanson

; REGISTRATION NUMBER: 33,215

; REFERENCE/DOCKET NUMBER: NEX21

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (303) 793-3333

; TELEFAX: (303) 793-3433

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 11

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-08-303-362A-6

## Query Match

Best Local Similarity 67.6%; Score 48; DB 1; Length 11;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11

Db 1 RPKPQQFFGLM 11

## RESULT 21

US-08-462-859A-1

; Sequence 1, Application US/08462859A

; Patent No. 5652092

; GENERAL INFORMATION:

; APPLICANT: Jacobsen, J. S.

; APPLICANT: Vitek, M. P.

; TITLE OF INVENTION: No. 5652092el Amyloid Precursor and Method of

; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation

; TITLE OF INVENTION: of B-Amyloid Peptide

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: American Cyanamid Company

; STREET: One Cyanamid Plaza

; CITY: Wayne

; STATE: New Jersey

; COUNTRY: United States

; ZIP: 07470-8426

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/462,859A

; FILING DATE: 05-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Barnhard, Elizabeth M.

; REGISTRATION NUMBER: 31,088

; REFERENCE/DOCKET NUMBER: 31,844-04

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (201)831-3246

; TELEFAX: (201)831-3305

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 11 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-462-859A-1

Query Match 67.6%; Score 48; DB 1; Length 11;

Best Local Similarity 81.8%; Pred. No. 0.1;

Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11

Db 1 RPKPQQFFGLM 11

## RESULT 22

US-08-123-659A-1

; Sequence 1, Application US/08123659A

; Patent No. 5656477

; GENERAL INFORMATION:

; APPLICANT: Jacobsen, J. S.

; APPLICANT: Vitek, M. P.

; TITLE OF INVENTION: No. 5656477el Amyloid Precursor and Method of

; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation

; TITLE OF INVENTION: of B-Amyloid Peptide

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Anne Rosenblum

; STREET: 163 Delaware Avenue, Suite 212

; CITY: Delmar

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 12054

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/123,659A  
; FILING DATE: 20-SEP-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rosenblum, Anne M.  
; REGISTRATION NUMBER: 30,419  
; REFERENCE/DOCKET NUMBER: 31,844-01  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (518)475-0611  
; TELEFAX: (518)475-0619  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-123-659A-1

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQQFFGLM 11

RESULT 23  
US-08-462-415-2  
; Sequence 2, Application US/08462415  
; Patent No. 5670499  
; GENERAL INFORMATION:  
; APPLICANT: Cho, Sung Y.  
; APPLICANT: Crowell, Thomas A.  
; APPLICANT: Gitter, Bruce D.  
; APPLICANT: Hipskind, Philip A.  
; APPLICANT: Howbert, Jeffrey J.  
; APPLICANT: Krushinski, Joseph H.  
; APPLICANT: Lobb, Karen L.  
; APPLICANT: Muehl, Brian S.  
; APPLICANT: Nixon, James A.  
; TITLE OF INVENTION: HETEROCYCLIC TACHYKININ RECEPTOR  
; TITLE OF INVENTION: ANTAGONISTS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center/Patent Division  
; CITY: Indianapolis  
; STATE: IN  
; COUNTRY: US  
; ZIP: 46285

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462,415  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gaylo, Paul J.  
; REGISTRATION NUMBER: 36,808  
; REFERENCE/DOCKET NUMBER: X8849B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-0756  
; TELEFAX: 317-276-3861  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid

; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-462-415-2

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQQFFGLM 11

RESULT 24  
US-08-463-874-2  
; Sequence 2, Application US/08463874  
; Patent No. 5684033  
; GENERAL INFORMATION:  
; APPLICANT: Cho, Sung Y.  
; APPLICANT: Crowell, Thomas A.  
; APPLICANT: Gitter, Bruce D.  
; APPLICANT: Hipskind, Philip A.  
; APPLICANT: Howbert, Jeffrey J.  
; APPLICANT: Krushinski, Joseph H.  
; APPLICANT: Lobb, Karen L.  
; APPLICANT: Muehl, Brian S.  
; APPLICANT: Nixon, James A.  
; TITLE OF INVENTION: NON-PEPTIDE TACHYKININ RECEPTOR  
; TITLE OF INVENTION: ANTAGONISTS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center/Patent Division  
; CITY: Indianapolis  
; STATE: IN  
; COUNTRY: US  
; ZIP: 46285  
COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,874  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gaylo, Paul J.  
; REGISTRATION NUMBER: 36,808  
; REFERENCE/DOCKET NUMBER: X8849C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 317-276-0756  
; TELEFAX: 317-276-3861  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-463-874-2

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQQFFGLM 11

```
RESULT 25
US-08-464-247A-1
; Sequence 1, Application US/08464247A
; Patent No. 5693478
; GENERAL INFORMATION:
; APPLICANT: Jacobsen, J. S.
; TITLE OF INVENTION: No. 5693478el Amyloid Precursor and Method of
; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
; TITLE OF INVENTION: of B-Amyloid Peptide
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: American Cyanamid Company
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,247A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 31,844-03
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-683-2158
; TELEFAX: 201-683-4117
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-247A-1

Query Match 67.68; Score 48; DB 1; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.1;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11
| | | | | | | |
Db 1 RPKPQQFFGLM 11

RESULT 26
US-08-464-248A-1
; Sequence 1, Application US/08464248A
; Patent No. 5703209
; GENERAL INFORMATION:
; APPLICANT: Jacobsen, J. S.
; TITLE OF INVENTION: No. 5703209el Amyloid Precursor and Method of
; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation
; TITLE OF INVENTION: of B-Amyloid Peptide
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07470-8426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,135
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/955,380
; FILING DATE: 01-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jarkovsky, Issac
; REGISTRATION NUMBER: 22,713
; REFERENCE/DOCKET NUMBER: 7754-003-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,248A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 31,844-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)831-3246
; TELEFAX: (201)831-3305
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-464-248A-1

Query Match 67.68; Score 48; DB 1; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.1;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11
| | | | | | | |
Db 1 RPKPQQFFGLM 11

RESULT 27
US-08-444-135-2
; Sequence 2, Application US/08444135
; Patent No. 5723575
; GENERAL INFORMATION:
; APPLICANT: Gilon, Chaim
; APPLICANT: Zeligler, Zvi
; TITLE OF INVENTION: Backbone Cyclic Peptides, Processes For
; TITLE OF INVENTION: Their Preparation and Pharmaceutical Compositions
; TITLE OF INVENTION: Containing Them
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,135
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/955,380
; FILING DATE: 01-OCT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jarkovsky, Issac
; REGISTRATION NUMBER: 22,713
; REFERENCE/DOCKET NUMBER: 7754-003-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 869-8864/9741
; TELEX: 66141 PENNIE
```

; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-444-135-2

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
Db 1 RPKPQQFFGLM 11

RESULT 28  
US-318-391-2  
; Sequence 2, Application US/08318391  
; Patent No. 574482  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Marlene L.  
; APPLICANT: Johnson, Kirk W.  
; APPLICANT: Phebus, Lee A.  
; TITLE OF INVENTION: USE OF A SEROTONIN AGONIST IN  
; TITLE OF INVENTION: COMBINATION WITH A TACHIKININ RECEPTOR ANTAGONIST IN THE  
; TITLE OF INVENTION: TREATMENT OR PREVENTION OF MIGRAINE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: United States of America  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/318,391  
; FILING DATE:  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gaylo, Paul J.  
; REGISTRATION NUMBER: 36,808  
; REFERENCE/DOCKET NUMBER: X-9664  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317) 276-0756  
; TELEFAX: (317) 276-3861  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-318-391-2

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
Db 1 RPKPQQFFGLM 11

RESULT 29  
US-08-796-598-11  
; Sequence 11, Application US/08796598  
; Patent No. 5827659  
; GENERAL INFORMATION:  
; APPLICANT: PATTERSON, DALE H.  
; APPLICANT: TARR, GEORGE E.  
; TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING  
; TITLE OF INVENTION: POLYMERS USING MASS SPECTROMETRY.  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Administrator - Testa, Hurwitz &  
; ADDRESSEE: Thibeault  
; STREET: High Street Tower, 125 High Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/796,598  
; FILING DATE: 07-FEB-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/446,055  
; FILING DATE: 19-MAY-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FLYNN Esq., Kerry A.  
; REGISTRATION NUMBER: 33,693  
; REFERENCE/DOCKET NUMBER: SYP-115  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-796-598-11

Query Match 67.6%; Score 48; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
Db 1 RPKPQQFFGLM 11

RESULT 30  
US-08-447-175A-11  
; Sequence 11, Application US/08447175A  
; Patent No. 5869240  
; GENERAL INFORMATION:  
; APPLICANT: PATTERSON, DALE H.  
; TITLE OF INVENTION: METHODS AND APPARATUS FOR SEQUENCING  
; TITLE OF INVENTION: POLYMERS WITH A STATISTICAL CERTAINTY USING MASS  
; TITLE OF INVENTION: SPECTROMETRY.  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patent Administrator - Testa, Hurwitz &  
; ADDRESSEE: Thibeault, LLP  
; STREET: High Street Tower, 125 High Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
Db 1 RPKPQQFFGLM 11



; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,175A  
; FILING DATE: 19-MAY-1995  
; CLASSIFICATION: 422  
; ATTORNEY/AGENT INFORMATION:  
; NAME: RAUSCHENBACH, Kurt  
; REGISTRATION NUMBER: 40,137  
; REFERENCE/DOCKET NUMBER: SYP-114  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 248-7000  
; TELEFAX: (617) 248-7100  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-447-175A-11

Query Match 67.68; Score 48; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
Db 1 RPKPQQFFGLM 11

RESULT 31  
US-07-737-371E-21  
; Sequence 21, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 21:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 9...9  
; OTHER INFORMATION: where Xaa at position 9 is D-Ala  
; US-07-737-371E-21

Query Match 67.68; Score 48; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11  
Db 1 RPKPQQFFGLM 11

RESULT 32  
US-07-737-371E-22  
; Sequence 22, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 22:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 9...9  
; OTHER INFORMATION: where Xaa at position 9 is Sar  
; US-07-737-371E-22

Query Match 67.68; Score 48; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RPKPQQWFWM 11

Db 1 RPKPQQFFXLM 11  
|||||:| |

## RESULT 33

US-07-737-371E-24  
; Sequence 24, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 9...9  
; OTHER INFORMATION: where Xaa at position 9 is D-Pro

Db 1 RPKPQQFFXLM 11  
|||||:| |

## RESULT 34

US-07-737-371E-27  
; Sequence 27, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 9...9  
; OTHER INFORMATION: where Xaa at position 9 is Me-Gly

Db 1 RPKPQQFFXLM 11  
|||||:| |

## RESULT 35

US-07-737-371E-65  
; Sequence 65, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990

Db 1 RPKPQQFFXLM 11  
|||||:| |

## RESULT 36

US-07-737-371E-66  
; Sequence 66, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990

Db 1 RPKPQQFFXLM 11  
|||||:| |

## RESULT 37

US-07-737-371E-67  
; Sequence 67, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990

; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 9...9  
; OTHER INFORMATION: where Xaa at position 9 is Me-Gly  
US-07-737-371E-27

Query Match 67.6%; Score 48; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFFXLM 11  
|||||:| |  
Db 1 RPKPQQFFXLM 11  
|||||:| |

## RESULT 38

US-07-737-371E-68  
; Sequence 68, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990

QY 1 RPKPQQFFXLM 11  
|||||:| |  
Db 1 RPKPQQFFXLM 11  
|||||:| |

## RESULT 39

US-07-737-371E-69  
; Sequence 69, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990

QY 1 RPKPQQFFXLM 11  
|||||:| |  
Db 1 RPKPQQFFXLM 11  
|||||:| |

ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00108/028002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
LOCATION: 9...9  
OTHER INFORMATION: where Xaa at position 9 is Me-Gly  
US-07-737-371E-65

Query Match 67.6%; Score 48; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQQFWFLM 11  
| | | | | | | | | | | | | |  
Db 1 RPKPQQQFFXLM 11

## RESULT 36

US-07-737-371E-77  
Sequence 77, Application US/07737371E  
Patent No. 5876948

GENERAL INFORMATION:  
APPLICANT: Yankner, Bruce A.  
TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
NEUROTOXIN INHIBITORS (AS AMENDED)  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/737,371E  
FILING DATE: 29-JUL-1991  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/559,172  
FILING DATE: 27-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00108/028002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154

INFORMATION FOR SEQ ID NO: 77:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-737-371E-77

Query Match 67.6%; Score 48; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQQFWFLM 11  
| | | | | | | | | | | | | |  
Db 1 RPKPQQQFFGLM 11

## RESULT 37

US-08-848-766A-1  
Sequence 1, Application US/08848766A  
Patent No. 5932551

GENERAL INFORMATION:  
APPLICANT: Caldwell, Charles G.  
APPLICANT: Chapman, Kevin T.  
APPLICANT: Durette, Philippe L.  
APPLICANT: Esser, Craig K.  
APPLICANT: Hagmann, William K.  
APPLICANT: Kopka, Ihor E.  
APPLICANT: Polo, Scott A.  
APPLICANT: Sahoo, Soumya P.  
TITLE OF INVENTION: SUBSTITUTED N-CARBOXYALKYLPEPTIDYL  
DERIVATIVES AS ANTIDEGENERATIVE ACTIVE AGENTS  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Merck & Co., Inc.  
STREET: P.O. Box 2000, 126 E. Lincoln Ave.  
CITY: Rahway  
STATE: NJ  
COUNTRY: USA  
ZIP: 07065-0900

COMPUTER READABLE FORM:  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/848,766A  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/873,905  
FILING DATE: 24-APR-1992

ATTORNEY/AGENT INFORMATION:  
NAME: Panzer, Curtis C  
REGISTRATION NUMBER: 33,752  
REFERENCE/DOCKET NUMBER: 183551A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-3199  
TELEFAX: 908-594-4720  
TELEX:

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-848-766A-1

Query Match 67.6%; Score 48; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQQFWFLM 11  
| | | | | | | | | | | | | |  
Db 1 RPKPQQQFFGLM 11

## RESULT 38

US-08-927-128-17  
; Sequence 17, Application US/08927128  
; Patent No. 6127150  
; GENERAL INFORMATION:  
; APPLICANT: Coolidge, Thomas  
; APPLICANT: Wagner, Fred  
; APPLICANT: ven Heeke, Gino  
; APPLICANT: Schuster, Sheldon  
; APPLICANT: Stout, Jay  
; APPLICANT: Wylie, Dwane  
; TITLE OF INVENTION: PURIFICATION DIRECTED CLOSING OF PEPTIDES  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Merchant & Gould  
; STREET: 3100 No. 6127150west Center, 90 S. 7th Street  
; CITY: Minneapolis  
; STATE: MN  
; COUNTRY: U.S.A.  
; ZIP: 55402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 1.5  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/927,128  
; FILING DATE: 05-SEP-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/680,004  
; FILING DATE: 15-JUL-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carter, Charles G  
; REGISTRATION NUMBER: 35,093  
; REFERENCE/DOCKET NUMBER: 8648.20USD1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 612/332-5300  
; TELEFAX: 612/332-9081  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; US-08-927-128-17

Query Match 67.6%; Score 48; DB 3; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:11  
Db 1 RPKPQQFFGLM 11

RESULT 39  
US-08-257-966-2  
; Sequence 2, Application US/08257966  
; Patent No. 6175013  
; GENERAL INFORMATION:  
; APPLICANT: Hipskind, Philip A.  
; APPLICANT: Howbert, James J.  
; APPLICANT: Muehl, Brian S.  
; TITLE OF INVENTION: IMIDAZOLINYL TACHYKININ RECEPTOR  
; ANTAGONISTS  
; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center/1104  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: United States of America  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/257,966  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gaylo, Paul J.  
; REGISTRATION NUMBER: 36,808  
; REFERENCE/DOCKET NUMBER: X-9197  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317) 276-0756  
; TELEFAX: (317) 276-3861  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-257-966-2

Query Match 67.6%; Score 48; DB 4; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:11  
Db 1 RPKPQQFFGLM 11

RESULT 40  
PCT-US95-05600-23  
; Sequence 23, Application PC/TUS9505600  
; GENERAL INFORMATION:  
; APPLICANT: GOLD, LARRY  
; APPLICANT: NIEUWLANDT, DAN  
; APPLICANT: WECKER, MATTHEW  
; APPLICANT: SCHNEIDER, DANIEL J.  
; APPLICANT: FEIGON, JULI  
; APPLICANT: ALLEN, PATRICK  
; APPLICANT: SULLENGER, BRUCE A.  
; APPLICANT: DOUDNA, JENNIFER, A.  
; TITLE OF INVENTION: HIGH-AFFINITY LIGANDS OF  
; INSULIN RECEPTOR ANTIBODIES, TACHYKININ SUBSTANCE  
; P. HIV INTEGRASE AND HIV-1 REVERSE TRANSCRIPTASE  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Swanson & Bratschun, L.L.C.  
; STREET: 8400 E. Prentice Avenue, Suite 200  
; CITY: Englewood  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MG  
; MEDIUM TYPE: storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05600

; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/238,863  
; FILING DATE: 06-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/248,632  
; FILING DATE: 24-MAY-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/303,362  
; FILING DATE: 09-SEPTEMBER-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/361,795  
; FILING DATE: 21-DECEMBER-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/117,991  
; FILING DATE: 08-SEPTEMBER-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/931,473  
; FILING DATE: 17-AUGUST-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/964,624  
; FILING DATE: 21-OCTOBER-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/536,428  
; FILING DATE: 11-JUNE-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/714,131  
; FILING DATE: 10-JUNE-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/536,428  
; FILING DATE: 11-JUNE-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barry J. Swanson  
; REGISTRATION NUMBER: 33,215  
; REFERENCE/DOCKET NUMBER: NEX17/PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 793-3333  
; TELEFAX: (303) 793-3433  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; PCT-US95-05600-23

Query Match 67.6%; Score 48; DB 5; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
Db 1 RPKPOQFFGLM 11

RESULT 41  
5441935-1  
; Patent No. 5441935  
; APPLICANT: Rozengurt, Enrique; Zachary, Ian; Woll, Penella  
; TITLE OF INVENTION: ROTH FACTOR RECEPTORS  
; NUMBER OF SEQUENCES: 10  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/939,587  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 814,064  
; FILING DATE: 23-DEC-1991  
; APPLICATION NUMBER: 411,536  
; FILING DATE: 29-NOV-1989  
; SEQ ID NO: 1;

; LENGTH: 11  
5441935-1

Query Match 67.6%; Score 48; DB 6; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
Db 1 RPKPOQFFGLM 11

RESULT 42  
5441935-8  
; Patent No. 5441935  
; APPLICANT: Rozengurt, Enrique; Zachary, Ian; Woll, Penella  
; TITLE OF INVENTION: ROTH FACTOR RECEPTORS  
; NUMBER OF SEQUENCES:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/939,587  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 814,064  
; FILING DATE: 23-DEC-1991  
; APPLICATION NUMBER: 411,536  
; FILING DATE: 29-NOV-1989  
; SEQ ID NO: 8;  
; LENGTH: 11  
5441935-8

Query Match 67.6%; Score 48; DB 6; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
Db 1 RPKPOQFFGLM 11

RESULT 43  
US-08-441-591-7  
; Sequence 7, Application US/08441591  
; Patent No. 5637682  
; GENERAL INFORMATION:  
; APPLICANT: NIEUWLANDT, D., GOLD, L. AND WECKER, M.  
; TITLE OF INVENTION: HIGH-AFFINITY  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIGANDS  
; TITLE OF INVENTION: TO THE TACHYKININ  
; TITLE OF INVENTION: SUBSTANCE P  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Swanson & Bratschun, L.L.C.  
; STREET: 8400 E. Prentice Avenue, Suite 200  
; CITY: Englewood  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/441,591  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/303,362  
; FILING DATE: 9-SEPTEMBER-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/714,131

; FILING DATE: 10-JUNE-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/931,473  
; FILING DATE: 17-AUGUST-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/117,991  
; FILING DATE: 8-SEPTEMBER 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/536,428  
; FILING DATE: 11-JUNE-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/964,624  
; FILING DATE: 21-OCTOBER-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barry J. Swanson  
; REGISTRATION NUMBER: 33,215  
; REFERENCE/DOCKET NUMBER: NEX21/C  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 793-3333  
; TELEFAX: (303) 793-3433  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-441-591-7

Query Match 67.6%; Score 48; DB 1; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.11;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
Db 1 RPKPQQFFGLM 11

RESULT 44  
US-08-303-362A-7  
; Sequence 7, Application US/08303362A  
; Patent No. 5648214  
; GENERAL INFORMATION:  
; APPLICANT: NIEWLANDT, D., GOLD, L. AND WECKER, M.  
; TITLE OF INVENTION: HIGH-AFFINITY  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIGANDS  
; TITLE OF INVENTION: TO THE TACHYKININ  
; TITLE OF INVENTION: SUBSTANCE P  
; NUMBER OF SEQUENCES: 66  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Swanson & Bratschun, L.L.C.  
; STREET: 8400 E. Prentice Avenue, Suite 200  
; CITY: Englewood  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MG storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/303,362A  
; FILING DATE: 9-SEPTEMBER-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/714,131  
; FILING DATE: 10-JUNE-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/931,473  
; FILING DATE: 17-AUGUST-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/117,991

; FILING DATE: 8-SEPTEMBER 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/536,428  
; FILING DATE: 11-JUNE-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/964,624  
; FILING DATE: 21-OCTOBER-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Barry J. Swanson  
; REGISTRATION NUMBER: 33,215  
; REFERENCE/DOCKET NUMBER: NEX21  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (303) 793-3333  
; TELEFAX: (303) 793-3433  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-303-362A-7

Query Match 67.6%; Score 48; DB 1; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.11;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
Db 1 RPKPQQFFGLM 11

RESULT 45  
US-08-505-250-27  
; Sequence 27, Application US/08505250  
; Patent No. 6183983  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Haruya  
; APPLICANT: Yamamoto, Keiji  
; APPLICANT: Suzuki, Kokichi  
; APPLICANT: Ikeda, Masahiro  
; APPLICANT: Sakagami, Masahiro  
; APPLICANT: Taniguchi, Makoto  
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD  
; FILE REFERENCE: 110-511  
; CURRENT APPLICATION NUMBER: US/08/505,250  
; CURRENT FILING DATE: 1995-11-29  
; EARLIER APPLICATION NUMBER: PCT/JP95/00298  
; EARLIER FILING DATE: 1995-02-27  
; EARLIER APPLICATION NUMBER: JP 198187/94  
; EARLIER FILING DATE: 1994-08-23  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 27  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: peptide  
US-08-505-250-27

Query Match 67.6%; Score 48; DB 4; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.11;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
Db 2 RPKPQQFFGLM 12

RESULT 46

US-08-505-250-53  
; Sequence 53, Application US/08505250  
; Patent No. 6183983  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Haruya  
; APPLICANT: Yamamoto, Keiji  
; APPLICANT: Suzuki, Kokichi  
; APPLICANT: Ikeda, Masahiro  
; APPLICANT: Sakagami, Masahiro  
; APPLICANT: Taniguchi, Makoto  
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD  
; FILE REFERENCE: 110-511  
; CURRENT APPLICATION NUMBER: US/08/505,250  
; EARLIER FILING DATE: 1995-11-29  
; EARLIER FILING DATE: 1995-02-27  
; EARLIER APPLICATION NUMBER: JP 198187/94  
; EARLIER FILING DATE: 1994-08-23  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 53  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; US-08-505-250-53

Query Match 67.6%; Score 48; DB 4; Length 12;  
Best Local Similarity 81.6%; Pred. No. 0.11;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOQFWFLM 11  
Db 2 RPKPQOQFFGLM 12

RESULT 47  
PCT-US92-06532-4  
; Sequence 4, Application PC/TUS9206532  
; GENERAL INFORMATION:  
; APPLICANT: Krause, James E.  
; TITLE OF INVENTION: Human Substance P Receptor  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SD  
; STREET: 800 N. Lindbergh Blvd.  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: U.S.A  
; ZIP: 63167  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/06532  
; FILING DATE: 19920805  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyer, Scott J.  
; REGISTRATION NUMBER: 25,275  
; REFERENCE/DOCKET NUMBER: 07-24(776)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)694-3117  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 12 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single

; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 12  
; OTHER INFORMATION: /label= amide  
; PCT-US92-06532-4

Query Match 67.6%; Score 48; DB 5; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.11;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOQFWFLM 11  
Db 2 RPKPQOQFFGLM 12

RESULT 48  
PCT-US95-05600-24  
; Sequence 24, Application PC/TUS9505600  
; GENERAL INFORMATION:  
; APPLICANT: GOLD, LARRY  
; APPLICANT: NIEUWLANDT, DAN  
; APPLICANT: WECKER, MATTHEW  
; APPLICANT: SCHNEIDER, DANIEL J.  
; APPLICANT: FEIGON, JULI  
; APPLICANT: ALLEN, PATRICK  
; APPLICANT: SULLENGER, BRUCE A.  
; APPLICANT: DOUDNA, JENNIFER, A.  
; TITLE OF INVENTION: HIGH-AFFINITY LIGANDS OF  
; TITLE OF INVENTION: INSULIN RECEPTOR ANTIBODIES, TACHYKININ SUBSTANCE  
; TITLE OF INVENTION: P. HIV INTEGRASE AND HIV-1 REVERSE TRANSCRIPTASE  
; NUMBER OF SEQUENCES: 239  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Swanson & Bratschun, L.L.C.  
; STREET: 8400 E. Prentice Avenue, Suite 200  
; CITY: Englewood  
; STATE: Colorado  
; COUNTRY: USA  
; ZIP: 80111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MG  
; MEDIUM TYPE: storage  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/05600  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/238,863  
; FILING DATE: 06-MAY-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/248,632  
; FILING DATE: 24-MAY-1994  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/303,362  
; FILING DATE: 09-SEPTEMBER-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/361,795  
; FILING DATE: 21-DECEMBER-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/117,991  
; FILING DATE: 08-SEPTEMBER-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/931,473  
; FILING DATE: 17-AUGUST-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/964,624  
; FILING DATE: 21-OCTOBER-1992

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/536,428  
;; FILING DATE: 11-JUNE-1990  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/714,131  
;; FILING DATE: 10-JUNE-1991  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/536,428  
;; FILING DATE: 11-JUNE-1990  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Barry J. Swanson  
;; REGISTRATION NUMBER: 33,215  
;; REFERENCE/DOCKET NUMBER: NEX17/PCT  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (303) 793-3333  
;; TELEFAX: (303) 793-3433  
;; INFORMATION FOR SEQ ID NO: 24:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 12 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; PCT-US95-05600-24

Query Match 67.6%; Score 48; DB 5; Length 12;  
Best Local Similarity 81.8%; Pred. No. 0.11;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQQFFGLM 11

## RESULT 49

US-08-890-157A-2  
;; Sequence 2, Application US/08890157A  
;; Patent No. 6063758  
;; GENERAL INFORMATION:  
;; APPLICANT: Douglas A. Lappi and Ronald G. Wiley  
;; TITLE OF INVENTION: Substance P-Saporin (SP-SAP) Conjugates And  
;; NUMBER OF SEQUENCES: 4  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Cooper and Dunham LLP  
;; STREET: 1185 Avenue of the Americas  
;; CITY: New York  
;; STATE: NY  
;; COUNTRY: US  
;; ZIP: 10036  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/890,157A  
;; FILING DATE: 09-JUL-1997  
;; CLASSIFICATION: 514  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Phillips, Peter J.  
;; REGISTRATION NUMBER: 29,691  
;; REFERENCE/DOCKET NUMBER: 53984  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212)278-0400  
;; TELEFAX: (212)331-0526  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 20 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; US-08-890-157A-2

Query Match 67.6%; Score 48; DB 3; Length 20;  
Best Local Similarity 81.8%; Pred. No. 0.18;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|  
Db 10 RPKPQQFFGLM 20

## RESULT 50

US-08-505-250-50  
;; Sequence 50, Application US/08505250  
;; Patent No. 6183983  
;; GENERAL INFORMATION:  
;; APPLICANT: Sato, Haruya  
;; APPLICANT: Yamamoto, Keiji  
;; APPLICANT: Suzuki, Kokichi  
;; APPLICANT: Ikeda, Masahiro  
;; APPLICANT: Sakagami, Masahiro  
;; APPLICANT: Taniguchi, Makoto  
;; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD  
;; FILE REFERENCE: 110-511  
;; CURRENT APPLICATION NUMBER: US/08/505,250  
;; CURRENT FILING DATE: 1995-11-29  
;; EARLIER APPLICATION NUMBER: PCT/JP95/00298  
;; EARLIER FILING DATE: 1995-02-27  
;; EARLIER APPLICATION NUMBER: JP 198187/94  
;; EARLIER FILING DATE: 1994-08-23  
;; NUMBER OF SEQ ID NOS: 53  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 50  
;; LENGTH: 20  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; FEATURE:  
;; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
;; OTHER INFORMATION: peptide  
;; US-08-505-250-50

Query Match 67.6%; Score 48; DB 4; Length 20;  
Best Local Similarity 81.8%; Pred. No. 0.18;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|  
Db 2 RPKPQQFFGLM 12

## RESULT 51

5268359-5  
;; Patent No. 5268359  
;; APPLICANT: HARMAR, ANTHONY J.; PASCALL, JOHN; MCKEOWN, ANN  
;; TITLE OF INVENTION: HUMAN TACHYKININS AND THEIR PRECURSOR  
;; NUMBER OF SEQUENCES: 7  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/285,964  
;; FILING DATE: 03-JUN-1987  
;; SEQ ID NO: 51  
;; LENGTH: 126  
;; 5268359-5

Query Match 67.6%; Score 48; DB 6; Length 126;  
Best Local Similarity 81.8%; Pred. No. 1.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|  
Db 58 RPKPQQFFGLM 68



## RESULT 52

5268359-2  
;PATENT NO. 5268359  
; APPLICANT: HARMAR, ANTHONY J.; PASCALL, JOHN; MCKEOWN, ANN  
; TITLE OF INVENTION: HUMAN TACHYKININS AND THEIR PRECURSOR  
; NUMBER OF SEQUENCES: 7  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/285.964  
; FILING DATE: 03-JUN-1987  
; SEQ ID NO: 2:  
; LENGTH: 130  
5268359-2

Query Match 67.6%; Score 48; DB 6; Length 130;  
Best Local Similarity 81.8%; Pred. No. 1.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWFWM 11  
|||||:|  
Db 58 RPKPOQFFGLM 68

## RESULT 53

US-08-462-859A-9  
; Sequence 9, Application US/08462859A  
; Patent No. 5652092  
; GENERAL INFORMATION:  
; APPLICANT: Jacobsen, J. S.  
; APPLICANT: Vitek, M. P.  
; TITLE OF INVENTION: No. 5652092el Amyloid Precursor and Method of  
; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation  
; TITLE OF INVENTION: of B-Amyloid Peptide  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: American Cyanamid Company  
; STREET: One Cyanamid Plaza  
; CITY: Wayne  
; STATE: New Jersey  
; COUNTRY: United States  
; ZIP: 07470-8426

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462.859A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhard, Elizabeth M.  
REGISTRATION NUMBER: 31,088  
REFERENCE/DOCKET NUMBER: 31,844-04  
TELEPHONE: (201)831-3246  
TELEFAX: (201)831-3305  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-462-859A-9

Query Match 67.6%; Score 48; DB 1; Length 487;  
Best Local Similarity 81.8%; Pred. No. 4.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWFWM 11  
|||||:|

Db 362 RPKPOQFFGLM 372

## RESULT 54

US-08-123-659A-9  
; Sequence 9, Application US/08123659A  
; Patent No. 5656477  
; GENERAL INFORMATION:  
; APPLICANT: Jacobsen, J. S.  
; APPLICANT: Vitek, M. P.  
; TITLE OF INVENTION: No. 5656477el Amyloid Precursor and Method of  
; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation  
; TITLE OF INVENTION: of B-Amyloid Peptide  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Anne Rosenblum  
; STREET: 163 Delaware Avenue, Suite 212  
; CITY: Delmar  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 12054

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/123.659A  
FILING DATE: 20-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Rosenblum, Anne M.  
REGISTRATION NUMBER: 30,419  
REFERENCE/DOCKET NUMBER: 31,844-01  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (518)475-0611  
TELEFAX: (518)475-0619  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-123-659A-9

Query Match 67.6%; Score 48; DB 1; Length 487;  
Best Local Similarity 81.8%; Pred. No. 4.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWFWM 11  
|||||:|

Db 362 RPKPOQFFGLM 372

## RESULT 55

US-08-464-247A-9  
; Sequence 9, Application US/08464247A  
; Patent No. 5693478  
; GENERAL INFORMATION:  
; APPLICANT: Jacobsen, J. S.  
; APPLICANT: Vitek, M. P.  
; TITLE OF INVENTION: No. 5693478el Amyloid Precursor and Method of  
; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation  
; TITLE OF INVENTION: of B-Amyloid Peptide  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: American Cyanamid Company  
; STREET: One Campus Drive  
; CITY: Parsippany  
; STATE: New Jersey  
; COUNTRY: United States  
; ZIP: 07054

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,247A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhard, Elizabeth M.  
REGISTRATION NUMBER: 31,088  
REFERENCE/DOCKET NUMBER: 31,844-03  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-683-2158  
TELEFAX: 201-683-4117  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-247A-9

Query Match 67.6%; Score 48; DB 1; Length 487;  
Best Local Similarity 81.8%; Pred. No. 4.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 362 RPKPQQFFGLM 372

## RESULT 56

US-08-464-248A-9  
Sequence 9, Application US/08464248A  
Patent No. 5703209  
GENERAL INFORMATION:  
APPLICANT: Jacobsen, J. S.  
TITLE OF INVENTION: No. 5703209el Amyloid Precursor and Method of  
REGISTRATION NUMBER: 31,088  
REFERENCE/DOCKET NUMBER: 31,844-03  
TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation  
of B-Amyloid Peptide  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: American Cyanamid Company  
STREET: One Cyanamid Plaza  
CITY: Wayne  
STATE: New Jersey  
COUNTRY: United States  
ZIP: 07470-8426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,248A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhard, Elizabeth M.  
REGISTRATION NUMBER: 31,088  
REFERENCE/DOCKET NUMBER: 31,844-02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201)831-3246  
TELEFAX: (201)831-3305  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 487 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-464-248A-9

Query Match 67.6%; Score 48; DB 1; Length 487;  
Best Local Similarity 81.8%; Pred. No. 4.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 362 RPKPQQFFGLM 372

## RESULT 57

US-08-462-859A-7  
Sequence 7, Application US/08462859A  
Patent No. 5652092  
GENERAL INFORMATION:  
APPLICANT: Jacobsen, J. S.  
TITLE OF INVENTION: No. 5652092el Amyloid Precursor and Method of  
REGISTRATION NUMBER: 31,088  
REFERENCE/DOCKET NUMBER: 31,844-04  
TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation  
of B-Amyloid Peptide  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: American Cyanamid Company  
STREET: One Cyanamid Plaza  
CITY: Wayne  
STATE: New Jersey  
COUNTRY: United States  
ZIP: 07470-8426  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,859A  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Barnhard, Elizabeth M.  
REGISTRATION NUMBER: 31,088  
REFERENCE/DOCKET NUMBER: 31,844-04  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (201)831-3246  
TELEFAX: (201)831-3305  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 492 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-462-859A-7

Query Match 67.6%; Score 48; DB 1; Length 492;  
Best Local Similarity 81.8%; Pred. No. 4.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:| |  
Db 362 RPKPQQFFGLM 372

RESULT 58  
US-08-123-659A-7  
Sequence 7, Application US/08123659A  
Patent No. 5656477  
GENERAL INFORMATION:  
APPLICANT: Jacobsen, J. S.  
TITLE OF INVENTION: No. 5656477el Amyloid Precursor and Method of

;; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation  
;; TITLE OF INVENTION: of B-Amyloid Peptide  
;; NUMBER OF SEQUENCES: 19  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Anne Rosenblum  
;; STREET: 163 Delaware Avenue, Suite 212  
;; CITY: Delmar  
;; STATE: New York  
;; COUNTRY: U.S.A.  
;; ZIP: 12054  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA: US/08/123,659A  
;; FILING DATE: 20-SEP-1993  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Rosenblum, Anne M.  
;; REGISTRATION NUMBER: 30,419  
;; REFERENCE/DOCKET NUMBER: 31,844-01  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (518)475-0611  
;; TELEFAX: (518)475-0619  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 492 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-123-659A-7

Query Match 67.6%; Score 48; DB 1; Length 492;  
Best Local Similarity 81.8%; Pred. No. 4.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
Db 362 RPKPQQFFGLM 372

RESULT 59  
US-08-464-247A-7  
;; Sequence 7, Application US/08464247A  
;; Patent No. 5693478  
;; GENERAL INFORMATION:  
;; APPLICANT: Jacobsen, J. S.  
;; TITLE OF INVENTION: No. 5693478el Amyloid Precursor and Method of  
;; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation  
;; TITLE OF INVENTION: of B-Amyloid Peptide  
;; NUMBER OF SEQUENCES: 19  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: American Cyanamid Company  
;; STREET: One Campus Drive  
;; CITY: Parsippany  
;; STATE: New Jersey  
;; COUNTRY: United States  
;; ZIP: 07054  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA: US/08/464,247A  
;; FILING DATE: 05-JUN-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Barnhard, Elizabeth M.

;; REGISTRATION NUMBER: 31,088  
;; REFERENCE/DOCKET NUMBER: 31,844-03  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 201-683-2158  
;; TELEFAX: 201-683-4117  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 492 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-464-247A-7

Query Match 67.6%; Score 48; DB 1; Length 492;  
Best Local Similarity 81.8%; Pred. No. 4.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
Db 362 RPKPQQFFGLM 372

RESULT 60  
US-08-464-248A-7  
;; Sequence 7, Application US/08464248A  
;; Patent No. 5703209  
;; GENERAL INFORMATION:  
;; APPLICANT: Jacobsen, J. S.  
;; TITLE OF INVENTION: No. 5703209el Amyloid Precursor and Method of  
;; TITLE OF INVENTION: Using Same to Access Agents Which Down-Regulate Formation  
;; TITLE OF INVENTION: of B-Amyloid Peptide  
;; NUMBER OF SEQUENCES: 19  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: American Cyanamid Company  
;; STREET: One Cyanamid Plaza  
;; CITY: Wayne  
;; STATE: New Jersey  
;; COUNTRY: United States  
;; ZIP: 07470-8426  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/464,248A  
;; FILING DATE: 05-JUN-1995  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Barnhard, Elizabeth M.  
;; REGISTRATION NUMBER: 31,088  
;; REFERENCE/DOCKET NUMBER: 31,844-02  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (201)831-3246  
;; TELEFAX: (201)831-3305  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 492 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-464-248A-7

Query Match 67.6%; Score 48; DB 1; Length 492;  
Best Local Similarity 81.8%; Pred. No. 4.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
Db 362 RPKPQQFFGLM 372

```
RESULT 61
US-07-737-371E-20
; Sequence 20, Application US/07737371E
; Patent No. 5876948
; GENERAL INFORMATION:
; APPLICANT: Yankner, Bruce A.
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-737-371E-20

Query Match 66.2%; Score 47; DB 2; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.14;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11
Db 1 RRPQQFFALM 11

RESULT 62
US-07-737-371E-17
; Sequence 17, Application US/07737371E
; Patent No. 5876948
; GENERAL INFORMATION:
; APPLICANT: Yankner, Bruce A.
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-737-371E-17

Query Match 64.8%; Score 46; DB 2; Length 11;
Best Local Similarity 81.8%; Pred. No. 0.2;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11
Db 1 RRPQQMFGLM 11

RESULT 63
US-07-737-371E-23
; Sequence 23, Application US/07737371E
; Patent No. 5876948
; GENERAL INFORMATION:
; APPLICANT: Yankner, Bruce A.
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
```

INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-737-371E-23

Query Match 64.8%; Score 45; DB 2; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.2;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOWFWLM 11  
Db 1 RPKPQOWFFPLM 11

RESULT 64  
US-08-346-849-7  
; Sequence 7, Application US/08346849  
; Patent No. 5670483  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Shuguang  
; APPLICANT: Lockshin, Curtis  
; APPLICANT: Rich, Alexander  
; APPLICANT: Holmes, Todd  
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY  
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173-4799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/346,849  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,326  
; FILING DATE: 28 DECEMBER 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brook, David E.  
; REGISTRATION NUMBER: 22,592  
; REFERENCE/DOCKET NUMBER: MIT-6008  
; TELEPHONE: (617) 861-9540  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-346-849-7

Query Match 63.4%; Score 45; DB 1; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOW 7  
Db 1 RPKPQOW 7

RESULT 65  
US-08-293-284A-7  
; Sequence 7, Application US/08293284A  
; Patent No. 5955343  
; GENERAL INFORMATION:  
; APPLICANT: Holmes, Todd  
; APPLICANT: Zhang, Shuguang  
; APPLICANT: Rich, Alexander  
; APPLICANT: DiPersio, C. Michael  
; APPLICANT: Lockshin, Curtis  
; TITLE OF INVENTION: STABLE MACROSCOPIC MEMBRANES FORMED BY  
; TITLE OF INVENTION: SELF-ASSEMBLY OF AMPHIPHILIC PEPTIDES AND USES  
; NUMBER OF SEQUENCES: 64  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.  
; STREET: Two Militia Drive  
; CITY: Lexington  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02173-4799  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/293,284A  
; FILING DATE: 22-AUG-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/973,326  
; FILING DATE: 28-DEC-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brook, David E.  
; REGISTRATION NUMBER: 22,592  
; REFERENCE/DOCKET NUMBER: MIT-6008A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 861-6240  
; TELEFAX: (617) 861-9540  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-293-284A-7

Query Match 63.4%; Score 45; DB 2; Length 9;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOW 7  
Db 1 RPKPQOW 7

RESULT 66  
US-07-899-205-1  
; Sequence 1, Application US/07899205  
; Patent No. 5288730  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Raymond  
; APPLICANT: Teall, Martin R.  
; APPLICANT: Swain, Christopher J.  
; APPLICANT: Williams, Brian J.  
; TITLE OF INVENTION: AZABICYCLIC COMPOUNDS PHARMACEUTICAL  
; TITLE OF INVENTION: COMPOSITIONS CONTAINING THEM AND THEIR USE IN THERAPY  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Merck & Co., Inc.  
STREET: 126 E. Lincoln Avenue  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/899,205  
FILING DATE: 19920616  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Polk, Manfred  
REGISTRATION NUMBER: 27,102  
REFERENCE/DOCKET NUMBER: T-1106  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-4285  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-07-899-205-1

Query Match 63.4%; Score 45; DB 1; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.29;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQEFFGLM 11

RESULT 67  
US-08-496-118-1  
Sequence 1, Application US/08496118  
Patent No. 5830854  
GENERAL INFORMATION:  
APPLICANT: Hargreaves, Richard J.  
TITLE OF INVENTION: THERAPEUTIC USE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Robert J. No. 5830854th  
STREET: 126 E. Lincoln Avenue - P. O. Box 2000  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,118  
FILING DATE: 27-JUNE-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: No. 5830854th, Robert J.  
REGISTRATION NUMBER: 27,366  
REFERENCE/DOCKET NUMBER: T-1213CA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-7262  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-496-118-1

Query Match 63.4%; Score 45; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.29;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQEFFGLM 11

RESULT 68  
US-07-737-371E-25  
Sequence 25, Application US/07737371E  
Patent No. 5876948  
GENERAL INFORMATION:  
APPLICANT: Yankner, Bruce A.  
TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
NEUROTOXIN INHIBITORS (AS AMENDED)  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/737,371E  
FILING DATE: 29-JUL-1991  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/559,172  
FILING DATE: 27-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00108/028002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-737-371E-25

Query Match 63.4%; Score 45; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.29;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
|||||:|  
Db 1 RPKPQEFFGLM 11

RESULT 69

PCT-US92-06532-1  
; Sequence 1, Application PC/TUS9206532  
; GENERAL INFORMATION:  
; APPLICANT: Krause, James E.  
; TITLE OF INVENTION: Human Substance P Receptor  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scott J. Meyer, Monsanto Co., A3SD  
; STREET: 800 N. Lindbergh Blvd.  
; CITY: St. Louis  
; STATE: Missouri  
; COUNTRY: U.S.A  
; ZIP: 63167  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/06532  
; FILING DATE: 19920805  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meyer, Scott J.  
; REGISTRATION NUMBER: 25,275  
; REFERENCE/DOCKET NUMBER: 07-24(776)A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (314)694-3117  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: AMINO ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 11  
; OTHER INFORMATION: /label= amide  
PCT-US92-06532-1

Query Match 63.4%; Score 45; DB 5; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.29;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 RPKPQWFILM 11  
| | | | | | | |  
Db 1 RPKPEQFFGLM 11

RESULT 70  
US-07-737-371E-9  
; Sequence 9, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 10 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-737-371E-9

Query Match 60.6%; Score 43; DB 2; Length 10;  
Best Local Similarity 80.0%; Pred. No. 0.52;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 2 RPKPQWFILM 11  
| | | | | | | |  
Db 1 RPKPQFFGLM 10

RESULT 71  
US-08-031-325A-26  
; Sequence 26, Application US/08031325A  
; Patent No. 5369094  
; GENERAL INFORMATION:  
; APPLICANT: Schally, Andrew V.  
; APPLICANT: Cai, Renzhi  
; TITLE OF INVENTION: POLYPEPTIDE BOMBESIN ANTAGONISTS  
; NUMBER OF SEQUENCES: 37  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OMRI M. BEHR, ESQ  
; STREET: 325 PIERSON AVENUE  
; CITY: EDISON  
; STATE: NEW JERSEY  
; COUNTRY: USA  
; ZIP: 08837  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/031,325A  
; FILING DATE: 15-MAR-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/619,747  
; FILING DATE: 29-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BEHR, OMRI M.  
; REGISTRATION NUMBER: 22,940  
; REFERENCE/DOCKET NUMBER: SHAL3.0-014  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908) 494-5240  
; TELEFAX: (908) 494-0428  
; TELEX: 511642 BEPATEDIN  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: 11  
;; OTHER INFORMATION: /note= "Res 11 - Met-NH2"  
US-08-031-325A-26

Query Match 60.6%; Score 43; DB 1; Length 11;  
Best Local Similarity 80.0%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
|||||:|  
Db 1 RPKPQQFFGL 10

RESULT 72  
US-07-737-371E-13  
; Sequence 13, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 13:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 11...11  
; OTHER INFORMATION: where xaa at position 11 is ethionine

Query Match 60.6%; Score 43; DB 2; Length 11;  
Best Local Similarity 80.0%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
|||||:|  
Db 1 RPKPQQFFGL 10

RESULT 73  
US-07-737-371E-14  
; Sequence 14, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 11...11  
; OTHER INFORMATION: where xaa at position 11 is Nle

Query Match 60.6%; Score 43; DB 2; Length 11;  
Best Local Similarity 80.0%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
|||||:|  
Db 1 RPKPQQFFGL 10

RESULT 74  
US-07-737-371E-16  
; Sequence 16, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA



```
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/737,371E
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-8906
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; LOCATION: 1...1
; OTHER INFORMATION: where xaa at position 1 is ethionine
US-07-737-371E-16

Query Match 60.6%; Score 43; DB 2; Length 11;
Best Local Similarity 80.0%; Pred. No. 0.58;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQOWFWLM 11
DB 2 PKPQOFFGLM 11

RESULT 75
US-07-737-371E-18
; Sequence 18, Application US/07737371E
; Patent No. 5876948
; GENERAL INFORMATION:
; APPLICANT: Yankner, Bruce A.
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/737,371E
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-8906
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; LOCATION: 9...9
; OTHER INFORMATION: where xaa at position 9 is Sar
; LOCATION: 11...11
; OTHER INFORMATION: where xaa at position 11 is Met(O2)
US-07-737-371E-55
```

```
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/737,371E
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-8906
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; LOCATION: 9...9
; OTHER INFORMATION: where xaa at position 9 is Sar
; LOCATION: 11...11
; OTHER INFORMATION: where xaa at position 11 is Met(O2)
US-07-737-371E-55

Query Match 60.6%; Score 43; DB 2; Length 11;
Best Local Similarity 80.0%; Pred. No. 0.58;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQOWFWLM 11
DB 2 PKPQOFFGLM 11

RESULT 76
US-07-737-371E-55
; Sequence 55, Application US/07737371E
; Patent No. 5876948
; GENERAL INFORMATION:
; APPLICANT: Yankner, Bruce A.
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/737,371E
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-8906
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; LOCATION: 9...9
; OTHER INFORMATION: where xaa at position 9 is Sar
; LOCATION: 11...11
; OTHER INFORMATION: where xaa at position 11 is Met(O2)
US-07-737-371E-55
```

Query Match 60.6%; Score 43; DB 2; Length 11;  
Best Local Similarity 80.0%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
| | | | | : |  
Db 1 RPKPQQFFXL 10

## RESULT 77

US-07-737-371E-61  
; Sequence 61, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 10...10  
; OTHER INFORMATION: where Xaa at location 10 is Me-Leu  
US-07-737-371E-61

Query Match 60.6%; Score 43; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 11  
| | | | | : |  
Db 1 RPKPQQFFGX 11

## RESULT 78

US-07-737-371E-63  
; Sequence 63, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.

; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 63:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 11...11  
; OTHER INFORMATION: where Xaa at position 11 is Me-Met  
US-07-737-371E-63

Query Match 60.6%; Score 43; DB 2; Length 11;  
Best Local Similarity 80.0%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
| | | | | : |  
Db 1 RPKPQQFFGL 10

## RESULT 79

US-07-737-371E-64  
; Sequence 64, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-737-371E-64

Query Match 60.6%; Score 43; DB 2; Length 11;  
Best Local Similarity 80.0%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOQFWL 10  
| | | | | | | |  
Db 1 RPKPOQFGL 10

RESULT 80  
US-07-737-371E-66  
; Sequence 66, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 66:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 10...10  
; OTHER INFORMATION: where Xaa at position 10 is Me-Leu  
US-07-737-371E-66

Query Match 60.6%; Score 43; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPOQFWLM 11  
| | | | | | | |  
Db 1 RPKPOQFFGX 11

RESULT 81  
US-08-747-137-34  
; Sequence 34, Application US/08747137  
; Patent No. 5945033  
; GENERAL INFORMATION:  
; APPLICANT: YEN, Richard C.K.  
; TITLE OF INVENTION: NON-CROSSLINKED PROTEIN PARTICLES FOR  
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC USE  
; NUMBER OF SEQUENCES: 184  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/747,137  
; FILING DATE: 12-NOV-1996  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/212,546  
; FILING DATE: 14-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/069,831  
; FILING DATE: 01-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/959,560  
; FILING DATE: 13-OCT-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/641,720  
; FILING DATE: 15-JAN-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 016197-00084005  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; FEATURE:  
; NAME/KEY: Modified-site  
; LOCATION: 11  
; OTHER INFORMATION: /product= "Met-Amide"  
US-08-747-137-34

Query Match 60.6%; Score 43; DB 2; Length 11;

Best Local Similarity 80.0%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRPQQQFWL 10  
| | | | | : | |  
Db 1 RRPQQQFFGL 10

## RESULT 82

US-08-505-250-34

; Sequence 34, Application US/08505250  
; Patent No. 6183983  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Haruya  
; APPLICANT: Yamamoto, Keiji  
; APPLICANT: Suzuki, Kokichi  
; APPLICANT: Ikeda, Masahiro  
; APPLICANT: Sakagami, Masahiro  
; APPLICANT: Taniguchi, Makoto  
; TITLE OF INVENTION: PROTEIN MODIFICATION METHOD  
; FILE REFERENCE: 110-511  
; CURRENT APPLICATION NUMBER: US/08/505,250  
; CURRENT FILING DATE: 1995-11-29  
; EARLIER APPLICATION NUMBER: PCT/JP95/00298  
; EARLIER FILING DATE: 1995-02-27  
; EARLIER APPLICATION NUMBER: JP 198187/94  
; EARLIER FILING DATE: 1994-08-23  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 34  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: peptide

US-08-505-250-34

Query Match 60.6%; Score 43; DB 4; Length 11;  
Best Local Similarity 80.0%; Pred. No. 0.58;  
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RRPQQQFWL 10  
| | | | | : | |  
Db 2 RRPQQQFFGL 11

## RESULT 83

US-08-701-846-2

; Sequence 2, Application US/08701846  
; Patent No. 5717069  
; GENERAL INFORMATION:  
; APPLICANT: Granados, Robert R.  
; TITLE OF INVENTION: DNA SEQUENCE CODING FOR A POLYPEPTIDE  
; TITLE OF INVENTION: WHICH ENHANCES VIRUS INFECTION OF HOST INSECTS  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Barnard, Brown & Michaels  
; STREET: 306 E. State St., Suite 220  
; CITY: Ithaca,  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 14850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/701,846  
; FILING DATE: 23-AUG-1996  
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/002,743  
; FILING DATE: 24-AUG-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michaels, Christopher A.  
; REGISTRATION NUMBER: 34,390  
; REFERENCE/DOCKET NUMBER: BTI-32  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (607)273-1711  
; TELEFAX: (607)273-2609  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 902 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-701-846-2

; Query Match 60.6%; Score 43; DB 1; Length 902;  
; Best Local Similarity 70.0%; Pred. No. 41;  
; Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQQFWLM 11

| | | | | : | |

Db 353 PYQIQWAWLM 362

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

| | | | | : | |

LOCATION: 5  
OTHER INFORMATION: /note= "Position 5 = Glu-NH2."  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 6  
OTHER INFORMATION: /note= "Position 6 = Glu-NH2."  
US-08-428-488-15

Query Match 59.2%; Score 42; DB 1; Length 11;  
Best Local Similarity 63.6%; Pred. No. 0.81;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
| | | | |  
Db 1 RPKPEEFGLM 11

## RESULT 85

US-07-737-371E-15  
Sequence 15, Application US/07737371E  
Patent No. 5876948  
GENERAL INFORMATION:  
APPLICANT: Yankner, Bruce A.  
TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/737,371E  
FILING DATE: 29-JUL-1991  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/559,172  
FILING DATE: 27-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00108/028002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-737-371E-15

Query Match 59.2%; Score 42; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.81;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
| | | | |  
Db 1 RPKPQQFFGLM 11

## RESULT 86

US-07-737-371E-19  
Sequence 19, Application US/07737371E  
Patent No. 5876948  
GENERAL INFORMATION:  
APPLICANT: Yankner, Bruce A.  
TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/737,371E  
FILING DATE: 29-JUL-1991  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/559,172  
FILING DATE: 27-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00108/028002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-737-371E-19

Query Match 59.2%; Score 42; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.81;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
| | | | |  
Db 1 RPPQQFFGLM 11

## RESULT 87

US-07-737-371E-29  
Sequence 29, Application US/07737371E  
Patent No. 5876948  
GENERAL INFORMATION:  
APPLICANT: Yankner, Bruce A.  
TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95

Query Match 59.2%; Score 42; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.81;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
| | | | |  
Db 1 RPKPQQFFGLM 11

```
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/737,371E
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; LOCATION: 5...5
; OTHER INFORMATION: where Xaa at position 5 is homocysteine
; LOCATION: 9...9
; OTHER INFORMATION: where Xaa at position 9 is homocysteine
US-07-737-371E-29

Query Match 59.2%; Score 42; DB 2; Length 11;
Best Local Similarity 72.7%; Pred. No. 0.81;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RRPQQWFWM 11
Db 1 RRPXPQFFXLM 11

RESULT 88
US-07-737-371E-57
; Sequence 57, Application US/07737371E
; Patent No. 5876948
; GENERAL INFORMATION:
; APPLICANT: Yankner, Bruce A.
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/737,371E
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-737-371E-57
```

```
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-737-371E-57

Query Match 57.7%; Score 41; DB 2; Length 8;
Best Local Similarity 87.5%; Pred. No. 1.6e+05;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPQQWF 8
Db 1 RRPQQFF 8

RESULT 89
US-07-737-371E-11
; Sequence 11, Application US/07737371E
; Patent No. 5876948
; GENERAL INFORMATION:
; APPLICANT: Yankner, Bruce A.
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/737,371E
; FILING DATE: 29-JUL-1991
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/559,172
; FILING DATE: 27-JUL-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00108/028002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-737-371E-11

Query Match 57.7%; Score 41; DB 2; Length 9;
Best Local Similarity 87.5%; Pred. No. 1.6e+05;
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RRPQQWF 8
Db 1 RRPQQFF 8
```

RESULT 90  
US-07-737-371E-26  
; Sequence 26, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 26:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 8...8  
; OTHER INFORMATION: where Xaa at position 8 is Me-Phe  
US-07-737-371E-26

Query Match 57.7%; Score 41; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 1.1;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
Db 1 RPKPQQFXGLM 11

RESULT 91  
US-07-737-371E-62  
; Sequence 62, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US

ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 62:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-737-371E-62

Query Match 57.7%; Score 41; DB 2; Length 11;  
Best Local Similarity 87.5%; Pred. No. 1.1;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 8  
Db 1 RPKPQQFF 8

RESULT 92  
US-07-737-371E-67  
; Sequence 67, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070

TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
LOCATION: 11...11  
OTHER INFORMATION: where xaa at position 11 is Me-Met  
US-07-737-371E-67

Query Match 57.7%; Score 41; DB 2; Length 11;  
Best Local Similarity 87.5%; Pred. No. 1.1;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWF 8  
|||||:1  
Db 1 RPKPQQFF 8

RESULT 93  
US-08-890-157A-1  
Sequence 1, Application US/08890157A  
Patent No. 6063758  
GENERAL INFORMATION:  
APPLICANT: Douglas A. Iappi and Ronald G. Wiley  
TITLE OF INVENTION: Substance P-Saporin (SP-SAP) Conjugates And  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper and Dunham LLP  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: US  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/890,157A  
FILING DATE: 09-JUL-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Phillips, Peter J.  
REGISTRATION NUMBER: 29,691  
REFERENCE/DOCKET NUMBER: 53984  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)278-0400  
TELEFAX: (212)391-0526  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 17 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-890-157A-1

Query Match 57.7%; Score 41; DB 3; Length 17;  
Best Local Similarity 87.5%; Pred. No. 1.8;  
Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQWF 8  
|||||:1  
Db 10 RPKPQQFF 17

RESULT 94  
US-07-737-371E-28  
Sequence 28, Application US/07737371E  
Patent No. 5876948  
GENERAL INFORMATION:  
APPLICANT: Yankner, Bruce A.  
TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
MOLECULES THAT INHIBIT NEUROTOXIN INHIBITORS (AS AMENDED)  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/737,371E  
FILING DATE: 29-JUL-1991  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/559,172  
FILING DATE: 27-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00108/028002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-737-371E-28

Query Match 56.3%; Score 40; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 1.6;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:11  
Db 1 RPKPQQFFCLM 11

RESULT 95  
US-09-060-726A-6  
Sequence 6, Application US/09060726A  
Patent No. 6225530  
GENERAL INFORMATION:  
APPLICANT: Weigel, Detlef  
APPLICANT: Salk Institute  
TITLE OF INVENTION: FLOWERING LOCUS T (FT) AND GENETICALLY  
MODIFIED PLANTS HAVING MODULATED FLOWER DEVELOPMENT  
FILE REFERENCE: SALKINS.026A  
CURRENT APPLICATION NUMBER: US/09/060,726A  
CURRENT FILING DATE: 1998-04-15  
NUMBER OF SEQ ID NOS: 13  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 104  
TYPE: PRT  
ORGANISM: Arabidopsis thaliana  
US-09-060-726A-6



Query Match 56.3%; Score 40; DB 4; Length 104;  
Best Local Similarity 40.0%; Pred. No. 14;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KPQQWFWM 11  
|:|:|:|:  
Db 44 PSPREHVV 53

## RESULT 96

US-09-236-080-2  
; Sequence 2, Application US/09236080  
; Patent No. 6242217

; GENERAL INFORMATION:  
; APPLICANT: Helen Meadows

; APPLICANT: Conrad Chapman  
; TITLE OF INVENTION: No. 6242217el Compounds

; FILE REFERENCE: GP30031  
; CURRENT APPLICATION NUMBER: US/09/236.080

; CURRENT FILING DATE: 1999-01-25  
; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 2

; LENGTH: 411  
; TYPE: PRT

; ORGANISM: Homo sapiens  
US-09-236-080-2

Query Match 56.3%; Score 40; DB 4; Length 411;  
Best Local Similarity 55.6%; Pred. No. 54;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQQWFWM 11  
|:|:|:|:  
Db 271 KPVVWFIL 279

## RESULT 97

US-09-236-080-6

; Sequence 6, Application US/09236080  
; Patent No. 6242217

; GENERAL INFORMATION:  
; APPLICANT: Helen Meadows

; APPLICANT: Conrad Chapman  
; TITLE OF INVENTION: No. 6242217el Compounds

; FILE REFERENCE: GP30031  
; CURRENT APPLICATION NUMBER: US/09/236.080

; CURRENT FILING DATE: 1999-01-25  
; NUMBER OF SEQ ID NOS: 6

; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 6

; LENGTH: 411  
; TYPE: PRT

; ORGANISM: Mus musculus  
US-09-236-080-6

Query Match 56.3%; Score 40; DB 4; Length 411;  
Best Local Similarity 55.6%; Pred. No. 54;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQQWFWM 11  
|:|:|:|:  
Db 271 KPVVWFIL 279

## RESULT 98

US-08-447-411-76

; Sequence 76, Application US/08447411  
; Patent No. 5773243

; GENERAL INFORMATION:  
; APPLICANT: FRITZINGER, DAVID C.  
; APPLICANT: BREDEHORST, REINHARD  
; APPLICANT: VOGEL, CARL-WILHELM  
; TITLE OF INVENTION: DNA ENCODING COBRA C3, CVF1, AND CVF2  
; NUMBER OF SEQUENCES: 81  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/447,411

; FILING DATE:  
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/043,747

; FILING DATE: 07-APR-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5773243man F.

; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 1126-101-0

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220

; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 76:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1333 amino acids  
; TYPE: amino acid

; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-447-411-76

Query Match 56.3%; Score 40; DB 1; Length 1333;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PQQWFWM 11  
|:|:|:|:  
Db 435 PESWLWM 442

## RESULT 99

US-08-662-227-34

; Sequence 34, Application US/08662227  
; Patent No. 592320

; GENERAL INFORMATION:  
; APPLICANT: VOGEL, CARL-WILHELM

; APPLICANT: BREDEHORST, REINHORST  
; APPLICANT: KOCK, MICHAEL

; APPLICANT: FRITZINGER, DAVID  
; TITLE OF INVENTION: RECOMBINANT PROCVF

; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.

; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY  
; CITY: ARLINGTON  
; STATE: VA

; COUNTRY: USA  
; ZIP: 22202

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/662,227  
; APPLICATION NUMBER: US/08/662,227  
; FILING DATE: 14-JUN-1996  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 1126-0107-0X  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1333 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-662-227-34

Query Match 56.3%; Score 40; DB 2; Length 1333;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQQFWLM 11  
I: I I I I  
Db 435 PESWLWM 442

RESULT 100  
US-09-017-947-34  
; Sequence 34, Application US/09017947  
; Patent No. 6303754  
; GENERAL INFORMATION:  
; APPLICANT: VOGEL, CARL-WILHELM  
; APPLICANT: BREDEHORST, REINHORST  
; APPLICANT: KOCK, MICHAEL  
; APPLICANT: FRITZINGER, DAVID  
; TITLE OF INVENTION: RECOMBINANT PROCVF  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: P.C.  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY  
; CITY: ARLINGTON  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/017,947  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/662,227  
; FILING DATE: 14-JUN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NORMAN F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 1126-0107-0X  
; TELEPHONE: 703-413-3000  
; TELEFAX: 703-413-2220  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 1333 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-017-947-34

Query Match 56.3%; Score 40; DB 4; Length 1333;  
Best Local Similarity 62.5%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQQFWLM 11  
I: I I I I  
Db 435 PESWLWM 442

RESULT 101  
US-08-462-413-2  
; Sequence 2, Application US/08462413  
; Patent No. 5530009  
; GENERAL INFORMATION:  
; APPLICANT: Cho, Sung Y.  
; APPLICANT: Copp, James D.  
; APPLICANT: Ginah, Francis O.  
; APPLICANT: Hansen, Guy J.  
; APPLICANT: Hipskind, Phillip A.  
; APPLICANT: Huff, Bret E.  
; APPLICANT: Martinelli, Michael J.  
; APPLICANT: Staszak, Michael A.  
; APPLICANT: Tharp-Taylor, Roger W.  
; TITLE OF INVENTION: PROCESS FOR PREPARING NON-PEPTIDYL  
; TITLE OF INVENTION: TACHYKININ RECEPTOR ANTAGONISTS  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Eli Lilly and Company  
; STREET: Lilly Corporate Center  
; CITY: Indianapolis  
; STATE: Indiana  
; COUNTRY: United States of America  
; ZIP: 46285  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/462,413  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/271,708  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gaylo, Paul J.  
; REGISTRATION NUMBER: 36,808  
; REFERENCE/DOCKET NUMBER: X-9475  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (317) 276-0756  
; TELEFAX: (317) 276-3861  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-462-413-2

Query Match 54.9%; Score 39; DB 1; Length 11;  
Best Local Similarity 72.7%; Pred. No. 2.3;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
||| ||| |||  
Db 1 RPKRQQPFGLM 11

## RESULT 102

US-07-737-371E-34  
; Sequence 34, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:  
; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07737371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 5...5  
; OTHER INFORMATION: where Xaa at position 5 is D-Cys  
US-07-737-371E-34

Query Match 54.9%; Score 39; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred No. 2.3;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
||||| ||| |||  
Db 1 RPKPQQCFGLM 11

## RESULT 103

PCT-US91-00899-6  
; Sequence 6, Application PC/TUS9100899  
; GENERAL INFORMATION:  
; APPLICANT: Lowe, John B.  
; TITLE OF INVENTION: Method and Products For the Synthesis of  
; TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,  
; TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned  
; TITLE OF INVENTION: Genetic Sequences That Determine These Structures  
; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/00899  
; FILING DATE: 19910214  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye Ph.D., Jean-Paul  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-021-55 PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-5940  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 299 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; FRAGMENT TYPE: C-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; TISSUE TYPE: Blood  
; CELL LINE: A431  
PCT-US91-00899-6

Query Match 54.9%; Score 39; DB 5; Length 299;  
Best Local Similarity 55.6%; Pred. No. 56;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 9  
||| ||| |||  
Db 64 RPKPQQRFWM 72

## RESULT 104

US-07-914-281-2  
; Sequence 2, Application US/07914281  
; Patent No. 5324663  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/914,281  
FILING DATE: 19920720  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 361 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-07-914-281-2

Query Match 54.9%; Score 39; DB 1; Length 361;  
Best Local Similarity 55.6%; Pred. No. 68;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
||: ||: ||  
Db 126 RPOGQRWI 134

RESULT 105  
US-08-393-246-2  
Sequence 2, Application US/08393246  
Patent No. 559590  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
RE  
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
RE  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
ADDRESS: P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202-  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,246  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 361 amino acids

TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-393-246-2  
Query Match 54.9%; Score 39; DB 1; Length 361;  
Best Local Similarity 55.6%; Pred. No. 68;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
||: ||: ||  
Db 126 RPOGQRWI 134

RESULT 106  
US-08-273-411-3  
Sequence 3, Application US/08273411  
Patent No. 5625124  
GENERAL INFORMATION:  
APPLICANT: Falk, Per  
APPLICANT: Gordon, Jeffrey I.  
TITLE OF INVENTION: Animal Model for Gastro-Intestinal  
Disease  
TITLE OF INVENTION: Disease  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: USA  
ZIP: 30309-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/273,411  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: WU106  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404) 815-6508  
TELEFAX: (404) 815-6555  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 361 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: Internal  
FEATURE:  
NAME/KEY: misc-feature  
LOCATION: 1..361  
OTHER INFORMATION:  
PUBLICATION INFORMATION:  
AUTHORS: Kukowska-Latallo, et al.  
JOURNAL: Genes & Development  
VOLUME: 4  
PAGES: 1288-1303  
DATE: 1990  
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 1 TO 361  
US-08-273-411-3

Query Match 54.9%; Score 39; DB 1; Length 361;

Best Local Similarity 55.6%; Pred. No. 68;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQOWFW 9  
||:|:|:|  
Db 126 RPQQRWIW 134

## RESULT 107

US-08-525-058A-2  
; Sequence 2, Application US/08525058A  
; Patent No. 5770420  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/525,058A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 361 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-525-058A-2

Query Match 54.9%; Score 39; DB 1; Length 361;  
Best Local Similarity 55.6%; Pred. No. 68;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQOWFW 9  
||:|:|:|  
Db 126 RPQQRWIW 134

## RESULT 108

US-08-696-731-2  
; Sequence 2, Application US/08696731  
; Patent No. 595347  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/696,731  
; FILING DATE: 14-AUG-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/393,246  
; FILING DATE:  
; APPLICATION NUMBER: US 08/220,433  
; FILING DATE: 30-MAR-1994  
; APPLICATION NUMBER: US 07/914,281  
; FILING DATE: 20-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 361 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-696-731-2

Query Match 54.9%; Score 39; DB 2; Length 361;  
Best Local Similarity 55.6%; Pred. No. 68;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQOWFW 9  
||:|:|:|  
Db 126 RPQQRWIW 134

## RESULT 109

US-09-042-531-2  
; Sequence 2, Application US/09042531  
; Patent No. 6268193  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/042.531  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/393.246  
;; FILING DATE:  
;; APPLICATION NUMBER: US 08/220.433  
;; FILING DATE: 30-MAR-1994  
;; APPLICATION NUMBER: US 07/914.281  
;; FILING DATE: 20-JUL-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Lavalleye, Jean-Paul M. P.  
;; REGISTRATION NUMBER: 31,451  
;; REFERENCE/DOCKET NUMBER: 2363-060-55  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703)521-4500  
;; TELEFAX: (703)486-2347  
;; TELEX: 248855 OPAT UR  
;; INFORMATION FOR SEQ ID NO: 2:  
;; LENGTH: 361 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
US-09-042-531-2

Query Match 54.9%; Score 39; DB 4; Length 361;  
Best Local Similarity 55.6%; Pred. No. 68;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFW 9  
||: ||: ||  
Db 126 RPOGQRWII 134

RESULT 110  
PCT-US91-00899-7  
;; Sequence 7, Application PC/TUS9100899  
;; GENERAL INFORMATION:  
;; APPLICANT: Lowe, John B.  
;; TITLE OF INVENTION: Method and Products For the Synthesis of  
;; TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,  
;; TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned  
;; TITLE OF INVENTION: Genetic Sequences That Determine These Structures  
;; NUMBER OF SEQUENCES: 16  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESSEE: P.C.  
;; STREET: 1755 Jefferson Davis Highway, Suite 400  
;; CITY: Arlington  
;; STATE: Virginia  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: PCT/US91/00899  
;; FILING DATE: 19910214  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Lavalleye Ph.D., Jean-Paul  
;; REGISTRATION NUMBER: 31,451  
;; REFERENCE/DOCKET NUMBER: 2363-021-55 PCT  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703)521-5940  
;; TELEFAX: (703)486-2347  
;; TELEX: 248855 OPAT UR  
;; INFORMATION FOR SEQ ID NO: 7:  
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 361 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: YES  
;; ORIGINAL SOURCE:  
;; ORGANISM: Homo sapiens  
;; TISSUE TYPE: Blood  
;; CELL LINE: A431  
PCT-US91-00899-7

Query Match 54.9%; Score 39; DB 5; Length 361;  
Best Local Similarity 55.6%; Pred. No. 68;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFW 9  
||: ||: ||  
Db 126 RPOGQRWII 134

RESULT 111  
US-07-914-281-11  
;; Sequence 11, Application US/07914281  
;; Patent No. 5324663  
;; GENERAL INFORMATION:  
;; APPLICANT: LOWE, JOHN B.  
;; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
;; TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
;; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
;; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
;; NUMBER OF SEQUENCES: 14  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
;; ADDRESSEE: P.C.  
;; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
;; CITY: Arlington  
;; STATE: Virginia  
;; COUNTRY: U.S.A.  
;; ZIP: 22202  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/914,281  
;; FILING DATE: 19920720  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Lavalleye, Jean-Paul M. P.  
;; REGISTRATION NUMBER: 31,451  
;; REFERENCE/DOCKET NUMBER: 2363-060-55  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703)521-4500  
;; TELEFAX: (703)486-2347  
;; TELEX: 248855 OPAT UR  
;; INFORMATION FOR SEQ ID NO: 11:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 374 amino acids  
;; TYPE: AMINO ACID  
;; TOPOLOGY: unknown  
;; MOLECULE TYPE: protein  
US-07-914-281-11

Query Match 54.9%; Score 39; DB 1; Length 374;  
Best Local Similarity 55.6%; Pred. No. 70;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFW 9  
||: ||: ||  
Db 139 RPOGQRWII 147



; FILING DATE: 30-MAR-1994  
; APPLICATION NUMBER: US 07/914,281  
; FILING DATE: 20-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 11:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 374 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-696-731-11

Query Match 54.9%; Score 39; DB 2; Length 374;  
Best Local Similarity 55.6%; Pred. No. 70;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
||: ||| |  
Db 139 RPOGQRW 147

RESULT 115  
US-09-042-531-11  
; Sequence 11, Application US/09042531  
; Patent No. 6268193  
; GENERAL INFORMATION:  
; APPLICANT: LONE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; ADDRESSEE: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/042,531  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,246  
; FILING DATE:  
; APPLICATION NUMBER: US 08/220,433  
; FILING DATE: 30-MAR-1994  
; APPLICATION NUMBER: US 07/914,281  
; FILING DATE: 20-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:  
; LENGTH: 374 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-09-042-531-11

Query Match 54.9%; Score 39; DB 4; Length 374;  
Best Local Similarity 55.6%; Pred. No. 70;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
||: ||| |  
Db 139 RPOGQRW 147

RESULT 116  
US-08-816-693A-51  
; Sequence 51, Application US/08816693A  
; Patent No. 5874241  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S  
; APPLICANT: Turek, Fred W  
; APPLICANT: Pinto, Lawrence H  
; TITLE OF INVENTION: Clock Gene and Gene Product  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/816,693A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5874241thrup, Thomas E  
; REGISTRATION NUMBER: 33,268  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 747 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-816-693A-51

Query Match 54.9%; Score 39; DB 2; Length 747;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWF 10  
| ||| |  
Db 316 KGQQW 323

RESULT 117  
US-08-885-291-51  
; Sequence 51, Application US/08885291A  
; Patent No. 6057125  
; GENERAL INFORMATION:



; APPLICANT: Takahashi, Joseph S.  
; APPLICANT: Turek, Fred W.  
; APPLICANT: Pinto, Lawrence H.  
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
; FILE REFERENCE: 0290-5  
; CURRENT APPLICATION NUMBER: US/08/885,291A  
; CURRENT FILING DATE: 1997-06-30  
; EARLIER APPLICATION NUMBER: 08/816,693  
; EARLIER FILING DATE: 1997-03-13  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 51  
; LENGTH: 747  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-08-885-291-51

Query Match 54.9%; Score 39; DB 3; Length 747;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQQWFNL 10  
| ||| ||  
Db 316 KGQQWIL 323

RESULT 118  
US-09-496-672-51  
; Sequence 51, Application US/09496672  
; Patent No. 6291429  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S.  
; APPLICANT: Turek, Fred W.  
; APPLICANT: Pinto, Lawrence H.  
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
; FILE REFERENCE: 0290-5  
; CURRENT APPLICATION NUMBER: US/09/496,672  
; CURRENT FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 08/885,291  
; PRIOR FILING DATE: 1997-06-30  
; PRIOR APPLICATION NUMBER: 08/816,693  
; PRIOR FILING DATE: 1997-03-13  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 51  
; LENGTH: 747  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-496-672-51

Query Match 54.9%; Score 39; DB 4; Length 747;  
Best Local Similarity 75.0%; Pred. No. 1.4e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQQWFNL 10  
| ||| ||  
Db 316 KGQQWIL 323

RESULT 119  
US-08-785-310A-8  
; Sequence 8, Application US/08785310A  
; Patent No. 5840532  
; GENERAL INFORMATION:  
; APPLICANT: McKnight, Steven L.  
; APPLICANT: Russell, David W.  
; TITLE OF INVENTION: Neuronal PAS Domain Protein  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/785,310A  
; APPLICATION NUMBER: US/08/785,310A  
; FILING DATE: 21-JAN-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A.  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UTSD:1226  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 816 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-785-310A-8

Query Match 54.9%; Score 39; DB 2; Length 816;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQQWFNL 10  
| ||| ||  
Db 319 KGQQWIL 326

RESULT 120  
US-08-816-693A-53  
; Sequence 53, Application US/08816693A  
; Patent No. 5874241  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S.  
; APPLICANT: Turek, Fred W.  
; APPLICANT: Pinto, Lawrence H.  
; TITLE OF INVENTION: Clock Gene and Gene Product  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/08/816,693A  
; APPLICATION NUMBER: US/08/816,693A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5874241thrup, Thomas E.  
; REGISTRATION NUMBER: 33,268  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5400  
; INFORMATION FOR SEQ ID NO: 53:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 816 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-816-693A-53

Query Match 54.9%; Score 39; DB 2; Length 816;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
| ||| ||  
DB 319 KGQQWIL 326

RESULT 121  
US-08-885-291-53  
; Sequence 53, Application US/08885291A  
; Patent No. 6057125  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S.  
; APPLICANT: Turek, Fred W.  
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
; FILE REFERENCE: 0290-5  
; CURRENT APPLICATION NUMBER: US/08/885,291A  
; CURRENT FILING DATE: 1997-06-30  
; EARLIER APPLICATION NUMBER: 08/816,693  
; EARLIER FILING DATE: 1997-03-13  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 53  
; LENGTH: 816  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-08-885-291-53

Query Match 54.9%; Score 39; DB 3; Length 816;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
| ||| ||  
DB 319 KGQQWIL 326

RESULT 122  
US-09-496-672-53  
; Sequence 53, Application US/09496672  
; Patent No. 6291429  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S.  
; APPLICANT: Turek, Fred W.  
; APPLICANT: Pinto, Lawrence H.  
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
; FILE REFERENCE: 0290-5  
; CURRENT APPLICATION NUMBER: US/09/496,672  
; CURRENT FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 08/885,291  
; PRIOR FILING DATE: 1997-06-30  
; PRIOR APPLICATION NUMBER: 08/816,693  
; PRIOR FILING DATE: 1997-03-13  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 53  
; LENGTH: 816  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-496-672-53

Query Match 54.9%; Score 39; DB 4; Length 816;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
| ||| ||  
DB 319 KGQQWIL 326

RESULT 123  
US-08-785-310A-7  
; Sequence 7, Application US/08785310A  
; Patent No. 5840532  
; GENERAL INFORMATION:  
; APPLICANT: McKnight, Steven L.  
; APPLICANT: Russell, David W.  
; TITLE OF INVENTION: Neuronal PAS Domain Protein  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/785,310A  
; FILING DATE: 21-JAN-1997  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UTSD:1226  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 824 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-785-310A-7

Query Match 54.9%; Score 39; DB 2; Length 824;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
| ||| ||  
DB 319 KGQQWIL 326

RESULT 124  
US-08-816-693A-52  
; Sequence 52, Application US/08816693A  
; Patent No. 5874241  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S  
; APPLICANT: Turek, Fred W  
; APPLICANT: Pinto, Lawrence H  
; TITLE OF INVENTION: Clock Gene and Gene Product  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Rocky, Milnamow & Katz

; STREET: Two Prudential Plaza, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/816,693A  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5874241thrup, Thomas E  
; REGISTRATION NUMBER: 33,268  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 52:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 824 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-08-816-693A-52

Query Match 54.9%; Score 39; DB 2; Length 824;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
| ||| ||  
Db 319 KGQOWIWL 326

## RESULT 125

US-08-885-291-52  
; Sequence 52, Application US/08885291A  
; Patent No. 6057125  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S.  
; APPLICANT: Turek, Fred W.  
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
; FILE REFERENCE: 0290-5  
; CURRENT APPLICATION NUMBER: US/08/885,291A  
; CURRENT FILING DATE: 1997-06-30  
; EARLIER APPLICATION NUMBER: 08/816,693  
; EARLIER FILING DATE: 1997-03-13  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 52  
; LENGTH: 824  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-08-885-291-52

Query Match 54.9%; Score 39; DB 3; Length 824;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
| ||| ||  
Db 319 KGQOWIWL 326

## RESULT 126

US-09-496-672-52

; Sequence 52, Application US/09496672  
; Patent No. 6291429  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S.  
; APPLICANT: Turek, Fred W.  
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
; FILE REFERENCE: 0290-5  
; CURRENT APPLICATION NUMBER: US/09/496,672  
; CURRENT FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 08/885,291  
; PRIOR FILING DATE: 1997-06-30  
; PRIOR APPLICATION NUMBER: 08/816,693  
; PRIOR FILING DATE: 1997-03-13  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 52  
; LENGTH: 824  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-09-496-672-52

Query Match 54.9%; Score 39; DB 4; Length 824;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
| ||| ||  
Db 319 KGQOWIWL 326

## RESULT 127

US-08-885-291-55  
; Sequence 55, Application US/08885291A  
; Patent No. 6057125  
; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S.  
; APPLICANT: Turek, Fred W.  
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
; FILE REFERENCE: 0290-5  
; CURRENT APPLICATION NUMBER: US/08/885,291A  
; CURRENT FILING DATE: 1997-06-30  
; EARLIER APPLICATION NUMBER: 08/816,693  
; EARLIER FILING DATE: 1997-03-13  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 55  
; LENGTH: 846  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-08-885-291-55

Query Match 54.9%; Score 39; DB 3; Length 846;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
| ||| ||  
Db 344 KGQOWIWL 351

## RESULT 128

US-09-107-847-2  
; Sequence 2, Application US/09107847  
; Patent No. 6100062  
; GENERAL INFORMATION:  
; APPLICANT: DUCKWORTH, DAVID  
; APPLICANT: MICHALOVICH, DAVID  
; TITLE OF INVENTION: NOVEL USE  
; NUMBER OF SEQUENCES: 2

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Ratner & Prestia  
;; STREET: P.O. Box 980  
;; CITY: Valley Forge  
;; STATE: PA  
;; COUNTRY: USA  
;; ZIP: 19482  
;;  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible  
;; OPERATING SYSTEM: DOS  
;; SOFTWARE: FastSeq for Windows Version 2.0  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/107,847  
;; FILING DATE: 30-JUN-1998  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: EP 97304996.8  
;; FILING DATE: 08-JUL-1997  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Prestia, Paul F  
;; REGISTRATION NUMBER: 23,031  
;; REFERENCE/DOCKET NUMBER: GH-30003  
;; TELEPHONE: 610-407-0700  
;; TELEFAX: 610-407-0701  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 846 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-09-107-847-2

Query Match 54.9%; Score 39; DB 3; Length 846;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
| | | | |  
Db 344 KGOQWIWL 351

RESULT 129  
US-09-496-672-55  
;; Sequence 55, Application US/09496672  
;; Patent No. 6291429  
;; GENERAL INFORMATION:  
;; APPLICANT: Takahashi, Joseph S.  
;; APPLICANT: Turek, Fred W.  
;; APPLICANT: Pinto, Lawrence H.  
;; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
;; FILE REFERENCE: 0290-5  
;; CURRENT APPLICATION NUMBER: US/09/496,672  
;; CURRENT FILING DATE: 2000-02-03  
;; PRIOR APPLICATION NUMBER: 08/885,291  
;; PRIOR FILING DATE: 1997-06-30  
;; PRIOR APPLICATION NUMBER: 08/816,693  
;; PRIOR FILING DATE: 1997-03-13  
;; NUMBER OF SEQ ID NOS: 55  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 55  
;; LENGTH: 846  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-496-672-55

Query Match 54.9%; Score 39; DB 4; Length 846;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 3 KPOQWFWL 10  
| | | | |  
Db 344 KGOQWIWL 351

RESULT 130  
US-08-816-693A-2  
;; Sequence 2, Application US/08816693A  
;; Patent No. 5874241  
;; GENERAL INFORMATION:  
;; APPLICANT: Takahashi, Joseph S  
;; APPLICANT: Turek, Fred W  
;; APPLICANT: Pinto, Lawrence H  
;; TITLE OF INVENTION: Clock Gene and Gene Product  
;; NUMBER OF SEQUENCES: 53  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Dressler, Rocky, Milnamow & Katz  
;; STREET: Two Prudential Plaza, Suite 4700  
;; CITY: Chicago  
;; STATE: Illinois  
;; COUNTRY: USA  
;; ZIP: 60601  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/816,693A  
;; FILING DATE:  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: No. 5874241thrup, Thomas E  
;; REGISTRATION NUMBER: 33,268  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 312-616-5400  
;; TELEFAX: 312-616-5460  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 855 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-816-693A-2

Query Match 54.9%; Score 39; DB 2; Length 855;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
| | | | |  
Db 344 KGOQWIWL 351

RESULT 131  
US-08-885-291-2  
;; Sequence 2, Application US/08885291A  
;; Patent No. 6057125  
;; GENERAL INFORMATION:  
;; APPLICANT: Takahashi, Joseph S.  
;; APPLICANT: Turek, Fred W.  
;; APPLICANT: Pinto, Lawrence H.  
;; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
;; FILE REFERENCE: 0290-5  
;; CURRENT APPLICATION NUMBER: US/08/885,291A  
;; CURRENT FILING DATE: 1997-06-30  
;; EARLIER APPLICATION NUMBER: 08/816,693  
;; EARLIER FILING DATE: 1997-03-13  
;; NUMBER OF SEQ ID NOS: 55  
;; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2  
; LENGTH: 855  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-08-885-291-2

Query Match 54.9%; Score 39; DB 3; Length 855;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQQWFNL 10  
| | | | |  
Db 344 KGQQWNL 351

## RESULT 132

US-09-496-672-2  
; Sequence 2, Application US/09496672  
; Patent No. 6291429

; GENERAL INFORMATION:  
; APPLICANT: Takahashi, Joseph S.  
; APPLICANT: Turek, Fred W.  
; APPLICANT: Pinto, Lawrence H.  
; TITLE OF INVENTION: CLOCK GENE AND GENE PRODUCT  
; FILE REFERENCE: 0290-5  
; CURRENT APPLICATION NUMBER: US/09/496,672  
; CURRENT FILING DATE: 2000-02-03  
; PRIOR APPLICATION NUMBER: 08/885,291  
; PRIOR FILING DATE: 1997-06-30  
; PRIOR APPLICATION NUMBER: 08/816,693  
; PRIOR FILING DATE: 1997-03-13  
; NUMBER OF SEQ ID NOS: 55  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 855  
; TYPE: PRT  
; ORGANISM: Mus musculus  
US-09-496-672-2

Query Match 54.9%; Score 39; DB 4; Length 855;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQQWFNL 10  
| | | | |  
Db 344 KGQQWNL 351

## RESULT 133

5441935-7

; Patent No. 5441935  
; APPLICANT: Rozenqurt, Enrique; Zachary, Ian; Woll, Penella  
; TITLE OF INVENTION: ROTH FACTOR RECEPTORS  
; NUMBER OF SEQUENCES:  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/939,587  
; FILING DATE: 03-SEP-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 814,064  
; FILING DATE: 23-DEC-1991  
; APPLICATION NUMBER: 411,536  
; FILING DATE: 29-NOV-1989  
; SEQ ID NO: 7:  
; LENGTH: 6  
5441935-7

Query Match 53.5%; Score 38; DB 6; Length 6;  
Best Local Similarity 83.3%; Pred. No. 1.6e+05;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 6 QWFWM 11  
: | | | | |  
Db 1 RWFWM 6

## RESULT 134

US-07-737-371E-54  
; Sequence 54, Application US/07737371E  
; Patent No. 5876948  
; GENERAL INFORMATION:

; APPLICANT: Yankner, Bruce A.  
; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
; NUMBER OF SEQUENCES: 77  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson, P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990

; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066

; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906

; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 54:  
; SEQUENCE CHARACTERISTICS:

; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:

; LOCATION: 7...7  
; OTHER INFORMATION: where xaa at position 7 is p-Chloro-Phe  
; LOCATION: 8...8  
; OTHER INFORMATION: where xaa at position 8 is p-Chloro-Phe  
US-07-737-371E-54

Query Match 53.5%; Score 38; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 3.2;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
| | | | |  
Db 1 RPKPQQXXGLM 11

## RESULT 135

US-08-700-651-14  
; Sequence 14, Application US/08700651B  
; Patent No. 6015882

; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.  
; APPLICANT: GUT, JIRI

;; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
;; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum

;; FILE REFERENCE: 480.19-4(HV)  
;; CURRENT APPLICATION NUMBER: US/08/700.651B  
;; CURRENT FILING DATE: 1997-08-14  
;; EARLIER APPLICATION NUMBER: 08/415.751  
;; EARLIER FILING DATE: 1995-04-03  
;; NUMBER OF SEQ ID NOS: 15  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 14  
;; LENGTH: 91  
;; TYPE: PRT  
;; ORGANISM: Cryptosporidium parvum  
;; FEATURE:  
;; OTHER INFORMATION: mutant/variant of SEQ ID NO:5  
US-08-700-651-14

Query Match 53.5%; Score 38; DB 3; Length 91;  
Best Local Similarity 62.5%; Pred. No. 25;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPOQWFWL 10  
||:| ||  
Db 83 KPDEWCWL 90

## RESULT 136

US-08-928-361B-19  
; Sequence 19, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/928,361B  
;; FILING DATE: 12-SEP-1997

;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 60/026,062  
;; FILING DATE: 13-SEP-1996  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Verny, Hana  
;; REGISTRATION NUMBER: 30,518  
;; REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 650-324-1677  
;; TELEFAX: 650-324-1678

;; INFORMATION FOR SEQ ID NO: 19:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 91 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-928-361B-19

Query Match 53.5%; Score 38; DB 3; Length 91;  
Best Local Similarity 62.5%; Pred. No. 25;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPOQWFWL 10  
||:| ||  
Db 83 KPDEWCWL 90

## RESULT 137

US-08-326-117B-8  
; Sequence 8, Application US/08326117B  
; Patent No. 5693491  
; GENERAL INFORMATION:  
; APPLICANT: BULLA, LEE A.  
; APPLICANT: JI, TAE  
; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS  
; TITLE OF INVENTION: TOXIN  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Ave. N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812

;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/326,117B  
;; FILING DATE: 19-OCT-1994  
;; CLASSIFICATION: 435  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: MILLMAN, ROBERT A.  
;; REGISTRATION NUMBER: 36,217  
;; REFERENCE/DOCKET NUMBER: 7112-0037.00

;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 887-1500  
;; TELEFAX: (202) 887-0763  
;; TELEX: 90-4030  
;; INFORMATION FOR SEQ ID NO: 8:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 113 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
US-08-326-117B-8

Query Match 53.5%; Score 38; DB 1; Length 113;  
Best Local Similarity 71.4%; Pred. No. 31;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PQQWFWL 10  
|:| | ||  
Db 57 PQQWFWL 63

## RESULT 138

US-08-982-129-8  
; Sequence 8, Application US/08982129  
; Patent No. 6007981  
; GENERAL INFORMATION:  
; APPLICANT: BULLA, LEE A.  
; APPLICANT: JI, TAE  
; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS  
; TITLE OF INVENTION: TOXIN  
; NUMBER OF SEQUENCES: 26

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Ave. N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/982,129  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/326,117  
; FILING DATE: 19-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MILLMAN, ROBERT A.  
; REGISTRATION NUMBER: 36,217  
; REFERENCE/DOCKET NUMBER: 7112-0037.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; TELEX: 90-4030  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 113 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-08-982-129-8

Query Match 53.5%; Score 38; DB 3; Length 113;  
Best Local Similarity 71.4%; Pred. No. 31;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 KPOQWFWL 10  
| | | | |  
Db 57 PNQWFWL 63

RESULT 139  
US-08-700-651-11  
; Sequence 11, Application US/08700651B  
; Patent No. 6015882  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum  
; FILE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-4(HV)  
; CURRENT APPLICATION NUMBER: US/08/700,651B  
; CURRENT FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 08/415,751  
; EARLIER FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 11  
; LENGTH: 124  
; TYPE: PRT  
; ORGANISM: Cryptosporidium parvum  
; FEATURE:  
; OTHER INFORMATION: mutant/variant of SEQ ID NO:5  
; US-08-700-651-11

Query Match 53.5%; Score 38; DB 3; Length 124;  
Best Local Similarity 62.5%; Pred. No. 34;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPOQWFWL 10  
| | | | |  
Db 116 KPDEWCWL 123

RESULT 140  
US-08-928-361B-16  
; Sequence 16, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,062  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verdy, Hana  
; REGISTRATION NUMBER: 30,518  
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-1677  
; TELEFAX: 650-324-1678  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 124 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-928-361B-16

Query Match 53.5%; Score 38; DB 3; Length 124;  
Best Local Similarity 62.5%; Pred. No. 34;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPOQWFWL 10  
| | | | |  
Db 116 KPDEWCWL 123

RESULT 141  
US-08-700-651-7  
; Sequence 7, Application US/08700651B  
; Patent No. 6015882  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.

APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum  
; TITLE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-4(HV)  
; CURRENT APPLICATION NUMBER: US/08/700,651B  
; CURRENT FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 08/415,751  
; EARLIER FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 128  
; TYPE: PRT  
; ORGANISM: Cryptosporidium parvum  
; FEATURE:  
; OTHER INFORMATION: mutant/variant of SEQ ID NO:5  
US-08-700-651-7

Query Match 53.5%; Score 38; DB 3; Length 128;  
Best Local Similarity 62.5%; Pred. No. 35;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
||:|:|  
Db 120 KPDEWCWL 127

RESULT 142  
US-08-928-361B-12  
; Sequence 12, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,062  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verny, Hana  
; REGISTRATION NUMBER: 30,518  
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-1677  
; TELEFAX: 650-324-1678  
; INFORMATION FOR SEQ ID NO: 12:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 128 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-928-361B-12

Query Match 53.5%; Score 38; DB 3; Length 128;  
Best Local Similarity 62.5%; Pred. No. 35;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
||:|:|  
Db 120 KPDEWCWL 127

RESULT 143  
US-08-700-651-8  
; Sequence 8, Application US/08700651B  
; Patent No. 6015882  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum  
; TITLE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-4(HV)  
; CURRENT APPLICATION NUMBER: US/08/700,651B  
; CURRENT FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 08/415,751  
; EARLIER FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 130  
; TYPE: PRT  
; ORGANISM: Cryptosporidium parvum  
; FEATURE:  
; OTHER INFORMATION: mutant/variant of SEQ ID NO:5  
US-08-700-651-8

Query Match 53.5%; Score 38; DB 3; Length 130;  
Best Local Similarity 62.5%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
||:|:|  
Db 122 KPDEWCWL 129

RESULT 144  
US-08-700-651-9  
; Sequence 9, Application US/08700651B  
; Patent No. 6015882  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum  
; TITLE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-4(HV)  
; CURRENT APPLICATION NUMBER: US/08/700,651B  
; CURRENT FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 08/415,751  
; EARLIER FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 130  
; TYPE: PRT  
; ORGANISM: Cryptosporidium parvum  
; FEATURE:



; OTHER INFORMATION: mutant/variant of SEQ ID NO:5  
US-08-700-651-9

Query Match 53.5%; Score 38; DB 3; Length 130;  
Best Local Similarity 62.5%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
|| : ||  
Db 122 KPDEWCWL 129

## RESULT 145

US-08-928-361B-13  
; Sequence 13, Application US/08928361B  
; Patent No. 6071518

## ; GENERAL INFORMATION:

; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PETERS, VERNY, JONES & BIKSA

; STREET: 385 Sherman Avenue, Suite 6

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306-1840

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/928,361B

; FILING DATE: 12-SEP-1997

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/026,062

; FILING DATE: 13-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: VERNY, HANA

; REGISTRATION NUMBER: 30,518

; REFERENCE/DOCKET NUMBER: 480.76-1(HV)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-324-1677

; TELEFAX: 650-324-1678

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 130 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-928-361B-13

Query Match 53.5%; Score 38; DB 3; Length 130;  
Best Local Similarity 62.5%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
|| : ||  
Db 122 KPDEWCWL 129

## RESULT 146

US-08-928-361B-14

; Sequence 14, Application US/08928361B

; Patent No. 6071518

## ; GENERAL INFORMATION:

; APPLICANT: Petersen, Carolyn

; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,

; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

; TITLE OF INVENTION: SPECIES INFECTIONS

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PETERS, VERNY, JONES & BIKSA

; STREET: 385 Sherman Avenue, Suite 6

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306-1840

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/928,361B

; FILING DATE: 12-SEP-1997

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/026,062

; FILING DATE: 13-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: VERNY, HANA

; REGISTRATION NUMBER: 30,518

; REFERENCE/DOCKET NUMBER: 480.76-1(HV)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-324-1677

; TELEFAX: 650-324-1678

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 130 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-928-361B-14

Query Match 53.5%; Score 38; DB 3; Length 130;  
Best Local Similarity 62.5%; Pred. No. 36;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQWFWL 10  
|| : ||  
Db 122 KPDEWCWL 129

## RESULT 147

US-08-700-651-10

; Sequence 10, Application US/08700651B

; Patent No. 6015882

; GENERAL INFORMATION:

; APPLICANT: PETERSEN, CAROLYN

; APPLICANT: LEECH, JAMES

; APPLICANT: NELSON, RICHARD, C.

; APPLICANT: GUT, JIRI

; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS

; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum

; TITLE OF INVENTION: INFECTIONS

; FILE REFERENCE: 480.19-4(HV)

; CURRENT APPLICATION NUMBER: US/08/700,651B

; CURRENT FILING DATE: 1997-08-14

; EARLIER APPLICATION NUMBER: 08/415,751

; EARLIER FILING DATE: 1995-04-03

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 10

; LENGTH: 138

; TYPE: PRT  
; ORGANISM: Cryptosporidium parvum  
; FEATURE:  
; OTHER INFORMATION: mutant/variant of SEQ ID NO:5  
US-08-700-651-10

Query Match 53.5%; Score 38; DB 3; Length 138;  
Best Local Similarity 62.5%; Pred. No. 38;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
||:|  
Db 130 KPDEWCWL 137

## RESULT 148

US-08-928-361B-15  
; Sequence 15, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,062  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verny, Hana  
; REGISTRATION NUMBER: 30,518  
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-1677  
; TELEFAX: 650-324-1678  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 138 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-928-361B-15

Query Match 53.5%; Score 38; DB 3; Length 138;  
Best Local Similarity 62.5%; Pred. No. 38;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
||:|  
Db 130 KPDEWCWL 137

## RESULT 149

US-08-928-361B-18  
; Sequence 18, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,062  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verny, Hana  
; REGISTRATION NUMBER: 30,518  
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-1677  
; TELEFAX: 650-324-1678  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 150 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-928-361B-18

Query Match 53.5%; Score 38; DB 3; Length 150;  
Best Local Similarity 62.5%; Pred. No. 41;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
||:|  
Db 142 KPDEWCWL 149

## RESULT 150

US-08-928-361B-9  
; Sequence 9, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928.361B  
FILING DATE: 12-SEP-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,062  
FILING DATE: 13-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: VERNY, Hana  
REGISTRATION NUMBER: 30,518  
REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
TELEPHONE: 650-324-1677  
TELEFAX: 650-324-1678  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 159 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-361B-9

Query Match 53.5%; Score 38; DB 3; Length 159;  
Best Local Similarity 62.5%; Pred. No. 43;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
||:||||  
Db 8 KPDEWCWL 15

RESULT 151  
US-08-928-361B-28  
; Sequence 28, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/928.361B  
FILING DATE: 12-SEP-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/026,062  
FILING DATE: 13-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: VERNY, Hana  
REGISTRATION NUMBER: 30,518  
REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-324-1677  
TELEFAX: 650-324-1678  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 159 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-361B-28

Query Match 53.5%; Score 38; DB 3; Length 159;  
Best Local Similarity 62.5%; Pred. No. 43;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
||:||||  
Db 8 KPDEWCWL 15

RESULT 152  
US-08-700-651-13  
; Sequence 13, Application US/08700651B  
; Patent No. 6015882  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF CRYPTOSPORIDIUM PARVUM  
; TITLE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-4(HV)  
; CURRENT APPLICATION NUMBER: US/08/700.651B  
; CURRENT FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 08/415,751  
; EARLIER FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 13  
; LENGTH: 162  
; TYPE: PRT  
; ORGANISM: Cryptosporidium parvum  
; FEATURE:  
; OTHER INFORMATION: mutant/variant of SEQ ID NO:5  
US-08-700-651-13

Query Match 53.5%; Score 38; DB 3; Length 162;  
Best Local Similarity 62.5%; Pred. No. 44;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQOWFWL 10  
||:||||  
Db 154 KPDEWCWL 161

RESULT 153  
US-08-700-651-12  
; Sequence 12, Application US/08700651B  
; Patent No. 6015882  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF CRYPTOSPORIDIUM PARVUM  
; TITLE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-4(HV)  
; CURRENT APPLICATION NUMBER: US/08/700.651B  
; CURRENT FILING DATE: 1997-08-14

; EARLIER APPLICATION NUMBER: 08/415,751  
; EARLIER FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 12  
; LENGTH: 175  
; TYPE: PRT  
; ORGANISM: Cryptosporidium parvum  
; FEATURE:  
; OTHER INFORMATION: mutant/variant of SEQ ID NO:5  
US-08-700-651-12

Query Match 53.5%; Score 38; DB 3; Length 175;  
Best Local Similarity 62.5%; Pred. No. 47;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQFWL 10  
|| : ||  
Db 167 KPDEWCWL 174

RESULT 154  
US-08-928-361B-17  
; Sequence 17, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B  
; FILING DATE: 12-SEP-1997

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,062  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verny, Hana  
; REGISTRATION NUMBER: 30,518  
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-1677  
; TELEFAX: 650-324-1678  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 175 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-928-361B-17  
Query Match 53.5%; Score 38; DB 3; Length 175;  
Best Local Similarity 62.5%; Pred. No. 47;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQFWL 10  
|| : ||  
Db 167 KPDEWCWL 174

RESULT 155  
US-08-700-651-15  
; Sequence 15, Application US/08700651B  
; Patent No. 6015882  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum  
; TITLE OF INVENTION: INFECTIONS  
; FILE REFERENCE: 480.19-4(HV)  
; CURRENT APPLICATION NUMBER: US/08/700,651B  
; CURRENT FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 08/415,751  
; EARLIER FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 15  
; LENGTH: 249  
; TYPE: PRT  
; ORGANISM: Cryptosporidium parvum  
; FEATURE:  
; OTHER INFORMATION: mutant/variant of SEQ ID NO:5  
US-08-700-651-15

Query Match 53.5%; Score 38; DB 3; Length 249;  
Best Local Similarity 62.5%; Pred. No. 67;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPOQFWL 10  
|| : ||  
Db 241 KPDEWCWL 248

RESULT 156  
US-08-928-361B-20  
; Sequence 20, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B  
; FILING DATE: 12-SEP-1997

CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,062  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verny, Hana  
; REGISTRATION NUMBER: 30,518  
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-1677  
; TELEFAX: 650-324-1678  
; INFORMATION FOR SEQ ID NO: 17:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 175 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein

US-08-928-361B-20  
Query Match 53.5%; Score 38; DB 3; Length 175;  
Best Local Similarity 62.5%; Pred. No. 47;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

NAME: Verny, Hana  
REGISTRATION NUMBER: 30,518  
REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-1677  
TELEFAX: 650-324-1678  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 249 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-928-361B-20

Query Match 53.5%; Score 38; DB 3; Length 249;  
Best Local Similarity 62.5%; Pred. No. 67;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQQFWL 10  
||:|:|  
Db 241 KPDEWCWL 248

RESULT 157  
US-09-055-095-4  
; Sequence 4, Application US/09055095  
; Patent No. 5945308  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Patterson, Chandra  
; APPLICANT: Corley, Neil C.  
; APPLICANT: Sather, Susan  
; TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/055,095  
; FILING DATE: Filed Herewith  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0500 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-855-0555  
; TELEFAX: 650-845-4166  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 270 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: GenBank  
; CLONE: 1902982  
US-09-055-095-4

Query Match 53.5%; Score 38; DB 2; Length 270;  
Best Local Similarity 62.5%; Pred. No. 72;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQWFW 9  
|:|:|  
Db 139 PCPQDWLW 146

RESULT 158  
US-08-809-494A-2  
; Sequence 2, Application US/08809494A  
; Patent No. 5962260  
; GENERAL INFORMATION:  
; APPLICANT: Sawamura, Tatsuya  
; APPLICANT: Masaki, Tomoo  
; TITLE OF INVENTION: Modified Low-Density Lipoprotein  
; TITLE OF INVENTION: Receptor  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel  
; STREET: 261 Madison Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10016-2391  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/809,494A  
; FILING DATE: 24-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-321705  
; FILING DATE: 30-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-214206  
; FILING DATE: 31-JUL-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldberg, Jules E.  
; REGISTRATION NUMBER: 24408  
; REFERENCE/DOCKET NUMBER: JG-YY-4363PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 986-4090  
; TELEFAX: 212 818-9479  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 270 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-809-494A-2

Query Match 53.5%; Score 38; DB 2; Length 270;  
Best Local Similarity 62.5%; Pred. No. 72;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQWFW 9  
|:|:|  
Db 139 PCPQDWLW 146

RESULT 159  
US-09-352-302-2  
; Sequence 2, Application US/09352302  
; Patent No. 6197937  
; GENERAL INFORMATION:  
; APPLICANT: Sawamura, Tatsuya

APPLICANT: Masaki, Tomoo  
TITLE OF INVENTION: Modified Low-Density Lipoprotein  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel  
STREET: 261 Madison Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10016-2391  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/352,302  
FILING DATE: 12-JUL-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-321705  
FILING DATE: 30-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-214206  
FILING DATE: 31-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldberg, Jules E  
REGISTRATION NUMBER: 24408  
REFERENCE/DOCKET NUMBER: JG-YY-4363PCT/D  
TELEPHONE: 212 986-4090  
TELEFAX: 212 818-9479  
INFORMATION FOR SEQ ID NO: 2:  
LENGTH: 270 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-352-302-2

Query Match 53.5%; Score 38; DB 4; Length 270;  
Best Local Similarity 62.5%; Pred. No. 72;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 2 PKPQQWF 9  
Db 139 PCPQDWLW 146

RESULT 160  
US-09-055-095-3  
Sequence 3, Application US/09055095  
Patent No. 5945308  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Patterson, Chandra  
APPLICANT: Corley, Neil C.  
APPLICANT: Sather, Susan  
TITLE OF INVENTION: HUMAN OXIDIZED LDL RECEPTOR  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/055,095  
FILING DATE: Filed Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0500 US  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
INFORMATION FOR SEQ ID NO: 3:  
LENGTH: 273 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GenBank  
CLONE: 1902984  
US-09-055-095-3

Query Match 53.5%; Score 38; DB 2; Length 273;  
Best Local Similarity 62.5%; Pred. No. 73;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQWF 9  
Db 143 PCPQDWLW 150

RESULT 161  
US-08-809-494A-4  
Sequence 4, Application US/08809494A  
Patent No. 5962260  
GENERAL INFORMATION:  
APPLICANT: Sawamura, Tatsuya  
APPLICANT: Masaki, Tomoo  
TITLE OF INVENTION: Modified Low-Density Lipoprotein  
TITLE OF INVENTION: Receptor  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel  
STREET: 261 Madison Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10016-2391  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/809,494A  
FILING DATE: 24-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-321705  
FILING DATE: 30-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-214206  
FILING DATE: 31-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldberg, Jules E  
REGISTRATION NUMBER: 24408  
REFERENCE/DOCKET NUMBER: JG-YY-4363PCT  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212 986-4090  
TELEFAX: 212 818-9479  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 273 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-809-494A-4

Query Match 53.5%; Score 38; DB 2; Length 273;  
Best Local Similarity 62.5%; Pred. No. 73;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQWF 9  
| | | |  
DB 142 PCPDWLW 149

RESULT 162  
US-08-809-494A-6  
; Sequence 6, Application US/08809494A  
; Patent No. 5962260  
; GENERAL INFORMATION:  
; APPLICANT: Sawamura, Tatsuya  
; APPLICANT: Masaki, Tomoo  
; TITLE OF INVENTION: Modified Low-Density Lipoprotein  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel  
; STREET: 261 Madison Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10016-2391  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/809,494A  
; FILING DATE: 24-MAR-1997  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-321705  
; FILING DATE: 30-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-214206  
; FILING DATE: 31-JUL-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldberg, Jules E  
; REGISTRATION NUMBER: 24408  
; REFERENCE/DOCKET NUMBER: JG-YY-4363PCT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 986-4090  
; TELEFAX: 212 818-9479  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 273 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-809-494A-6

Query Match 53.5%; Score 38; DB 2; Length 273;  
Best Local Similarity 62.5%; Pred. No. 73;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQWF 9

DB 143 PCPDWLW 150  
| | | |

RESULT 163  
US-09-352-302-4  
; Sequence 4, Application US/09352302  
; Patent No. 6197937  
; GENERAL INFORMATION:  
; APPLICANT: Sawamura, Tatsuya  
; APPLICANT: Masaki, Tomoo  
; TITLE OF INVENTION: Modified Low-Density Lipoprotein  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel  
; STREET: 261 Madison Avenue  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10016-2391  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/352,302  
; FILING DATE: 12-JUL-1999  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 6-321705  
; FILING DATE: 30-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 7-214206  
; FILING DATE: 31-JUL-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Goldberg, Jules E  
; REGISTRATION NUMBER: 24408  
; REFERENCE/DOCKET NUMBER: JG-YY-4363PCT/D  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 986-4090  
; TELEFAX: 212 818-9479  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 273 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-352-302-4

Query Match 53.5%; Score 38; DB 4; Length 273;  
Best Local Similarity 62.5%; Pred. No. 73;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQWF 9  
| | | |  
DB 142 PCPDWLW 149

RESULT 164  
US-09-352-302-6  
; Sequence 6, Application US/09352302  
; Patent No. 6197937  
; GENERAL INFORMATION:  
; APPLICANT: Sawamura, Tatsuya  
; APPLICANT: Masaki, Tomoo  
; TITLE OF INVENTION: Modified Low-Density Lipoprotein  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: McAulay Fisher Nissen Goldberg & Kiel

STREET: 261 Madison Avenue  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10016-2391  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/352.302  
FILING DATE: 12-JUL-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 6-321705  
FILING DATE: 30-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 7-214206  
FILING DATE: 31-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Goldberg, Jules E.  
REGISTRATION NUMBER: 24408  
REFERENCE/DOCKET NUMBER: JG-YY-4363PCT/D  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 986-4090  
TELEFAX: 212 818-9479  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 273 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-352-302-6

Query Match 53.5%; Score 38; DB 4; Length 273;  
Best Local Similarity 62.5%; Pred. No. 73;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQWF 9  
| | | |  
Db 143 PCPDWIW 150

RESULT 165  
PCT-US91-00899-14  
SEQUENCE 14, Application PC/TUS9100899  
GENERAL INFORMATION:  
APPLICANT: Lowe, John B.  
TITLE OF INVENTION: Method and Products For the Synthesis of  
TITLE OF INVENTION: Oligosaccharide Structures on Glycoproteins, Glycolipids,  
TITLE OF INVENTION: or as Free Molecules, and For the Isolation of Cloned  
TITLE OF INVENTION: Genetic Sequences That Determine These Structures  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/00899  
FILING DATE: 19910214  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye Ph.D., Jean-Paul

REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-021-55 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-5940  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 357 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: C-terminal  
PCT-US91-00899-14

Query Match 53.5%; Score 38; DB 5; Length 357;  
Best Local Similarity 50.0%; Pred. No. 95;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKQWF 10  
| | | |  
Db 110 RPKQWF 119

RESULT 166  
US-07-914-281-8  
SEQUENCE 8, Application US/07914281  
PATENT NO. 5324663  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
TITLE OF INVENTION: OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURES  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: P.C.  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/914,281  
FILING DATE: 19920720  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 405 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-07-914-281-8

Query Match 53.5%; Score 38; DB 1; Length 405;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;



Query Match 53.5%; Score 38; DB 1; Length 405;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
|||  
Db 161 RPPGQRWVM 170

RESULT 167  
US-08-393-246-8  
; Sequence 8, Application US/08393246  
; Patent No. 5595900  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOPOLIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,246  
; FILING DATE:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/393,246  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/220,433  
; FILING DATE: 30-MAR-1994  
; APPLICATION NUMBER: US 07/914,281  
; FILING DATE: 20-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 405 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-393-246-8

Query Match 53.5%; Score 38; DB 1; Length 405;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
|||  
Db 161 RPPGQRWVM 170

RESULT 168  
US-08-525-058A-8  
; Sequence 8, Application US/08525058A  
; Patent No. 5770420  
; GENERAL INFORMATION:  
; APPLICANT: LOWE, JOHN B.  
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
; TITLE OF INVENTION: GLYCOPOLIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
; TITLE OF INVENTION: OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTURE  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT, P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U.S.A.  
; ZIP: 22202  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/525,058A  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lavalleye, Jean-Paul M. P.  
; REGISTRATION NUMBER: 31,451  
; REFERENCE/DOCKET NUMBER: 2363-060-55  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)521-4500  
; TELEFAX: (703)486-2347  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 405 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
; US-08-525-058A-8

Query Match 53.5%; Score 38; DB 1; Length 405;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQFWL 10  
|||  
Db 161 RPPGQRWVM 170

RESULT 169  
US-08-483-151-4  
; Sequence 4, Application US/08483151  
; Patent No. 5858752  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian  
; APPLICANT: Holgersson, Jan  
; TITLE OF INVENTION: FUCOSYLTRANSFERASE GENES AND USES THEREOF  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/483,151  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lech, Karen F.  
; REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/278001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 405 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-483-151-4

Query Match 53.5%; Score 38; DB 2; Length 405;  
Best Local Similarity 40.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQOWFWL 10  
|||::|||  
Db 161 RPRKRWVM 170

RESULT 170  
US-08-696-731-8  
Sequence 8, Application US/08696731  
Patent No. 5953347  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/696,731  
FILING DATE: 14-AUG-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/393,246  
FILING DATE:  
APPLICATION NUMBER: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 405 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein

US-08-696-731-8

Query Match 53.5%; Score 38; DB 2; Length 405;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQOWFWL 10  
|||::|||  
Db 161 RPRKRWVM 170

RESULT 171  
US-09-042-531-8  
Sequence 8, Application US/09042531  
Patent No. 6288193  
GENERAL INFORMATION:  
APPLICANT: LOWE, JOHN B.  
TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS  
OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,  
TITLE OF INVENTION: GLYCOLIPIDS, OR AS FREE MOLECULES, AND FOR THE ISOLATION  
OF CLONED GENETIC SEQUENCES THAT DETERMINE THESE STRUCTU  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
STREET: 1755 Jefferson Davis Highway, Fourth Floor  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/042,531  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/393,246  
FILING DATE:  
APPLICATION NUMBER: US 08/220,433  
FILING DATE: 30-MAR-1994  
APPLICATION NUMBER: US 07/914,281  
FILING DATE: 20-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Lavalleye, Jean-Paul M. P.  
REGISTRATION NUMBER: 31,451  
REFERENCE/DOCKET NUMBER: 2363-060-55  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)521-4500  
TELEFAX: (703)486-2347  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 405 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-042-531-8

Query Match 53.5%; Score 38; DB 4; Length 405;  
Best Local Similarity 50.0%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQOWFWL 10  
|||::|||  
Db 161 RPRKRWVM 170

RESULT 172  
US-08-918-914-1  
; Sequence 1, Application US/08918914  
; Patent No. 5876963  
; GENERAL INFORMATION:  
; APPLICANT: Mitchell, Peter  
; APPLICANT: Hutchinson, Nancy  
; APPLICANT: Lawton, Michael  
; APPLICANT: Magna, Holly  
; APPLICANT: Vocum, Sue  
; APPLICANT: Murry, Lynn E.  
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Dr.  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/918,914  
; FILING DATE: Filed Herewith  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0369  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1184 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; LIBRARY: ???  
; CLONE: 422069  
US-08-918-914-1

Query Match 53.5%; Score 38; DB 2; Length 1184;  
Best Local Similarity 44.4%; Pred. No. 3e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWF 9  
Db 335 KRPDKYFW 343

RESULT 173  
US-08-996-083-3  
; Sequence 3, Application US/08996083A  
; Patent No. 6124095  
; GENERAL INFORMATION:  
; APPLICANT: Magna, Holly  
; APPLICANT: Schaffer, Paul  
; APPLICANT: Lawton, Michael  
; APPLICANT: Vocum, Sue  
; APPLICANT: Mitchell, Peter  
; APPLICANT: Hutchinson, Nancy  
; APPLICANT: Murry, Lynn E.  
; TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHORYLASE-2

; FILE REFERENCE: PF-0420 US  
; CURRENT APPLICATION NUMBER: US/08/996,083A  
; CURRENT FILING DATE: 1997-12-22  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 1184  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; OTHER INFORMATION: Incyte Clone No. 6124095: 422069  
; PUBLICATION INFORMATION:  
US-08-996-083-3

Query Match 53.5%; Score 38; DB 3; Length 1184;  
Best Local Similarity 44.4%; Pred. No. 3e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPOQWF 9  
Db 335 KRPDKYFW 343

RESULT 174  
US-08-326-117B-2  
; Sequence 2, Application US/08326117B  
; Patent No. 5693491  
; GENERAL INFORMATION:  
; APPLICANT: BULLA, LEE A.  
; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS  
; TITLE OF INVENTION: TOXIN  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Ave. N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/326,117B  
; FILING DATE: 19-OCT-1994  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MILLMAN, ROBERT A  
; REGISTRATION NUMBER: 36,217  
; REFERENCE/DOCKET NUMBER: 7112-0037.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; TELEX: 90-4030  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1528 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-326-117B-2

Query Match 53.5%; Score 38; DB 1; Length 1528;  
Best Local Similarity 71.4%; Pred. No. 3.9e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PQQWF 10

Db 233 PNQMWL 239  
| | | |

## RESULT 175

US-08-982-129-2  
; Sequence 2, Application US/08982129  
; Patent No. 6007981  
; GENERAL INFORMATION:  
; APPLICANT: BULLA, LEE A.  
; APPLICANT: JI, TAE  
; TITLE OF INVENTION: RECEPTOR FOR A BACILLUS THURINGIENSIS  
; TITLE OF INVENTION: TOXIN  
; NUMBER OF SEQUENCES: 26  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 2000 Pennsylvania Ave. N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20006-1812  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/982,129  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/326,117  
; FILING DATE: 19-OCT-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MILLMAN, ROBERT A  
; REGISTRATION NUMBER: 36,217  
; REFERENCE/DOCKET NUMBER: 7112-0037.00  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 887-1500  
; TELEFAX: (202) 887-0763  
; TELEX: 90-4030  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1528 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-982-129-2

Query Match 53.5%; Score 38; DB 3; Length 1528;  
Best Local Similarity 71.4%; Pred. No. 3.9e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQQWFWL 10  
| | | |

Db 233 PNQMWL 239

## RESULT 176

US-08-700-651-5  
; Sequence 5, Application US/08700651B  
; Patent No. 6015882  
; GENERAL INFORMATION:  
; APPLICANT: PETERSEN, CAROLYN  
; APPLICANT: LEECH, JAMES  
; APPLICANT: NELSON, RICHARD, C.  
; APPLICANT: GUT, JIRI  
; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS  
; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum  
; FILE REFERENCE: 480.19-4(HV)  
; CURRENT APPLICATION NUMBER: US/08/700,651B

; CURRENT FILING DATE: 1997-08-14  
; EARLIER APPLICATION NUMBER: 08/415,751  
; EARLIER FILING DATE: 1995-04-03  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 1721  
; TYPE: PRT  
; ORGANISM: Cryptosporidium parvum  
US-08-700-651-5

Query Match 53.5%; Score 38; DB 3; Length 1721;  
Best Local Similarity 62.5%; Pred. No. 4.4e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQQWFWL 10  
| | | |

Db 415 KPDEWCWL 422

## RESULT 177

US-08-928-361B-6  
; Sequence 6, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,062  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: VERNY, HANA  
; REGISTRATION NUMBER: 30,518  
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-1677  
; TELEFAX: 650-324-1678  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1721 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-928-361B-6

Query Match 53.5%; Score 38; DB 3; Length 1721;  
Best Local Similarity 62.5%; Pred. No. 4.4e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQQWFWL 10

Db 415 KPDEWCWL 422  
|| : || ||

## RESULT 178

US-08-928-361B-5

; Sequence 5, Application US/08928361B

; Patent No. 6071518

; GENERAL INFORMATION:

; APPLICANT: Petersen, Carolyn

; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,

; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM

; TITLE OF INVENTION: SPECIES INFECTIONS

; NUMBER OF SEQUENCES: 30

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PETERS, VERNY, JONES &amp; BIKSA

; STREET: 385 Sherman Avenue, Suite 6

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306-1840

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/928,361B

; FILING DATE: 12-SEP-1997

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/026,062

; FILING DATE: 13-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: Verny, Hana

; REGISTRATION NUMBER: 30,518

; REFERENCE/DOCKET NUMBER: 480,76-1(HV)

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-324-1677

; TELEFAX: 650-324-1678

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1837 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-928-361B-5

Query Match 53.5%; Score 38; DB 3; Length 1837;

Best Local Similarity 62.5%; Pred. No. 4.6e+02;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQOWFLM 10

|| : || ||

Db 532 KPDEWCWL 539

## RESULT 179

US-07-737-371E-3

; Sequence 3, Application US/07737371E

; Patent No. 5876948

; GENERAL INFORMATION:

; APPLICANT: Yankner, Bruce A.

; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY

; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)

; NUMBER OF SEQUENCES: 77

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish &amp; Richardson, P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/737,371E  
; FILING DATE: 29-JUL-1991  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/559,172  
; FILING DATE: 27-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00108/028002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-542-5070  
; TELEFAX: 617-542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 11 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; LOCATION: 1...1  
; OTHER INFORMATION: where xaa at position 1 is D-Arg  
; LOCATION: 7...7  
; OTHER INFORMATION: where xaa at position 7 is D-Trp  
; LOCATION: 9...9  
; OTHER INFORMATION: where xaa at position 9 is D-Trp  
US-07-737-371E-3

Query Match 52.1%; Score 37; DB 2; Length 11;

Best Local Similarity 70.0%; Pred. No. 4.6;

Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQOWFLM 11

||||| |:

Db 2 PKPQOWFLM 11

## RESULT 180

US-07-737-371E-30

; Sequence 30, Application US/07737371E

; Patent No. 5876948

; GENERAL INFORMATION:

; APPLICANT: Yankner, Bruce A.

; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY

; TITLE OF INVENTION: NEUROTOXIN INHIBITORS (AS AMENDED)

; NUMBER OF SEQUENCES: 77

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish &amp; Richardson, P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/737,371E

; FILING DATE: 29-JUL-1991

; CLASSIFICATION: 536

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/559,172  
;; FILING DATE: 27-JUL-1990  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Freeman, John W.  
;; REGISTRATION NUMBER: 29,066  
;; REFERENCE/DOCKET NUMBER: 00108/028002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 617-542-5070  
;; TELEFAX: 617-542-8906  
;; TELEX: 200154

;; INFORMATION FOR SEQ ID NO: 30:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids

;; TYPE: amino acid  
;; TOPOLOGY: linear

;; MOLECULE TYPE: protein  
;; FEATURE:

;; LOCATION: 5...5  
;; OTHER INFORMATION: where xaa at position 5 is homocysteine

;; LOCATION: 10...10  
;; OTHER INFORMATION: where xaa at position 10 is homocysteine

;; US-07-737-371E-30

Query Match 52.1%; Score 37; DB 2; Length 11;  
Best Local Similarity 63.6%; Pred. No. 4.6;  
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:|:|  
Db 1 RPKPQQFFGX 11

## RESULT 181

US-07-737-371E-32  
; Sequence 32, Application US/07737371E  
; Patent No. 5876948

;; GENERAL INFORMATION:  
;; APPLICANT: Yankner, Bruce A.

;; TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
;; NUMBER OF SEQUENCES: 77

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Fish & Richardson, P.C.

;; STREET: 225 Franklin Street  
;; CITY: Boston

;; STATE: MA  
;; COUNTRY: US

;; ZIP: 02110-2804  
;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Diskette  
;; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: Windows95  
;; SOFTWARE: FastSeq for Windows Version 2.0

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/07/737,371E

;; FILING DATE: 29-JUL-1991  
;; CLASSIFICATION: 536

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 07/559,172

;; FILING DATE: 27-JUL-1990  
;; ATTORNEY/AGENT INFORMATION:

;; NAME: Freeman, John W.  
;; REGISTRATION NUMBER: 29,066

;; REFERENCE/DOCKET NUMBER: 00108/028002  
;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: 617-542-5070  
;; TELEFAX: 617-542-8906

;; TELEX: 200154  
;; INFORMATION FOR SEQ ID NO: 32:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 11 amino acids

;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FEATURE:  
;; LOCATION: 5...5  
;; OTHER INFORMATION: where xaa at position 5 is homocysteine  
;; LOCATION: 11...11  
;; OTHER INFORMATION: where xaa at position 11 is homocysteine  
;; US-07-737-371E-32

Query Match 52.1%; Score 37; DB 2; Length 11;  
Best Local Similarity 70.0%; Pred. No. 4.6;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 10  
|||||:|:|  
Db 1 RPKPQQFFGL 10

## RESULT 182

US-08-468-514-11  
; Sequence 11, Application US/08468514  
; Patent No. 5576296

;; GENERAL INFORMATION:  
;; APPLICANT: Bartfal, Tamas

;; APPLICANT: Hofkfelt, Tomas  
;; APPLICANT: Langel, Ulo

;; APPLICANT: Ahren, Bo  
;; APPLICANT: Lindskog, Stefan

;; APPLICANT: Consolo, Silvana  
;; APPLICANT: Land, Tilt

;; APPLICANT: Wiesenfeld-Hallin, Zsuzsanna  
;; TITLE OF INVENTION: GALANIN ANTAGONIST

;; NUMBER OF SEQUENCES: 11  
;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: White & Case  
;; STREET: 1155 Avenue of the Americas

;; CITY: New York  
;; STATE: NY

;; COUNTRY: USA  
;; ZIP: 10036-2787

;; COMPUTER READABLE FORM: disk  
;; MEDIUM TYPE: Floppy disk

;; OPERATING SYSTEM: IBM PC compatible  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/468,514

;; FILING DATE:  
;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/146,139

;; FILING DATE: 12-NOV-1993  
;; APPLICATION NUMBER: PCT/SE92/00316

;; FILING DATE: 14-MAY-1992  
;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: SE 9101472-0  
;; FILING DATE: 15-MAY-1991

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Steiner Ph.D., Richard J.

;; REGISTRATION NUMBER: 35,372  
;; REFERENCE/DOCKET NUMBER: 1103326-074

;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 212-819-8783

;; TELEFAX: 212-354-8113  
;; INFORMATION FOR SEQ ID NO: 11:

;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 22 amino acids

;; TYPE: amino acid  
;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide  
;; HYPOTHETICAL: NO

```
;
;
; NAME/KEY: Modified-site
; LOCATION: 22
; OTHER INFORMATION: /note= "amide"
; US-08-468-514-11

Query Match 52.1%; Score 37; DB 1; Length 22;
Best Local Similarity 70.0%; Pred. No. 8.9;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 RPKQOWFWLM 11
Db 13 PPOQOFFGLM 22

RESULT 183
US-08-690-095-1
; Sequence 1, Application US/08690095
; Patent No. 5792648
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Au-Young, Janice
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN MACROPHAGE ANTIGEN
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: U.S.
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/113,789
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/690,095
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0110 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 272 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: MPHGN0T03
; CLONE: 513418
; US-09-113-789-1

Query Match 52.1%; Score 37; DB 3; Length 272;
Best Local Similarity 44.4%; Pred. No. 1e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKQOWFW 9
Db 131 KPCPRRWI 139

RESULT 185
US-09-028-934-35
; Sequence 35, Application US/09028934
; Patent No. 6117670
; GENERAL INFORMATION:
; APPLICANT: Ligon, James M.
; APPLICANT: Hill, Dwight S.
; APPLICANT: Lam, Steven T.
; APPLICANT: Hammer, Philip E.
; APPLICANT: van Pee, Karl-Heinz
; APPLICANT: Kirner, Sabine
; APPLICANT: Young, Thomas R.
; TITLE OF INVENTION: Pyrrolnitrin Biosynthesis Genes and Uses
; TITLE OF INVENTION: thereof
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6117670artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park

Query Match 52.1%; Score 37; DB 1; Length 272;
Best Local Similarity 44.4%; Pred. No. 1e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKQOWFW 9
Db 131 KPCPRRWI 139

RESULT 184
US-09-113-789-1
; Sequence 1, Application US/09113789
```

STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/028,934  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/729,214  
FILING DATE: 09-OCT-1996  
PRIOR APPLICATION DATA: US 08/258,261  
FILING DATE: 08-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC1506/CIP7  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 565 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-028-934-35

Query Match 52.1%; Score 37; DB 3; Length 565;  
Best Local Similarity 55.6%; Pred. No. 2.1e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
DB 227 KPGQRWRW 235

RESULT 186  
US-07-737-371E-60  
Sequence 60, Application US/07737371E  
Patent No. 5876948  
GENERAL INFORMATION:  
APPLICANT: Yankner, Bruce A.  
TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/737,371E  
FILING DATE: 29-JUL-1991  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/559,172  
FILING DATE: 27-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00108/028002  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
LOCATION: 2...2  
OTHER INFORMATION: where xaa at position 2 is D-Pro  
LOCATION: 7...7  
OTHER INFORMATION: where xaa at position 7 is D-Trp  
LOCATION: 9...9  
OTHER INFORMATION: where xaa at position 9 is D-Trp

REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00108/028002  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 9 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-737-371E-60

Query Match 50.7%; Score 36; DB 2; Length 9;  
Best Local Similarity 77.8%; Pred. No. 1.6e+05;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPOQFWLM 11  
DB 1 KPOQFGLM 9

RESULT 187  
US-07-737-371E-2  
Sequence 2, Application US/07737371E  
Patent No. 5876948  
GENERAL INFORMATION:  
APPLICANT: Yankner, Bruce A.  
TITLE OF INVENTION: SCREENING METHODS TO IDENTIFY  
NUMBER OF SEQUENCES: 77  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson, P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/737,371E  
FILING DATE: 29-JUL-1991  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/559,172  
FILING DATE: 27-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00108/028002  
TELEPHONE: 617-542-5070  
TELEFAX: 617-542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 11 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
LOCATION: 2...2  
OTHER INFORMATION: where xaa at position 2 is D-Pro  
LOCATION: 7...7  
OTHER INFORMATION: where xaa at position 7 is D-Trp  
LOCATION: 9...9  
OTHER INFORMATION: where xaa at position 9 is D-Trp



US-07-737-371E-2

Query Match 50.7%; Score 36; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 6.4;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQOWFWLM 11  
| | | | | | | |  
Db 1 RXPQOXFXLM 11

RESULT 188

US-07-712-828B-5

; Sequence 5, Application US/07712828B  
; Patent No. 5235039

; GENERAL INFORMATION:

; APPLICANT: Heath et al.

; TITLE OF INVENTION: Assay Method for Hydrolytic

; TITLE OF INVENTION: Enzymes

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Eli Lilly and Company

; STREET: Lilly Corporate Center

; CITY: Indianapolis

; STATE: IN.

; COUNTRY: U.S.A.

; ZIP: 46285

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 Mb storage

; COMPUTER: Macintosh

; OPERATING SYSTEM: Macintosh

; SOFTWARE: Microsoft Word

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07712.828B

; FILING DATE: 19010610

; CLASSIFICATION: 530

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 13 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

US-07-712-828B-5

Query Match 50.7%; Score 36; DB 1; Length 13;  
Best Local Similarity 63.6%; Pred. No. 7.6;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQOWFWLM 11  
| | | | | | | |  
Db 1 RRRPQOWFXLM 11

RESULT 189

US-09-082-089-3

; Sequence 3, Application US/09082089  
; Patent No. 6100060

; GENERAL INFORMATION:

; APPLICANT: BARNES, MICHAEL

; APPLICANT: TESTA, TANIA

; APPLICANT: KELSELL, DAVID

; TITLE OF INVENTION: No. 6100060e1 Compounds

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER &amp; PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082.089  
; FILING DATE: 20-MAY-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9710737.9  
; FILING DATE: 23-MAY-1997  
; APPLICATION NUMBER: GB 9803981.1  
; FILING DATE: 25-FEB-1998  
; APPLICATION NUMBER: GB 9804007.4  
; FILING DATE: 25-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PRESTIA, PAUL F.  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GH-30166  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 359 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-082-089-3

Query Match 50.7%; Score 36; DB 3; Length 359;  
Best Local Similarity 54.5%; Pred. No. 1.9e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 RPK--PQOWFW 9  
| | | | | | | |  
Db 148 RPKDLPDOWLM 158

RESULT 190

US-09-082-089-5

; Sequence 5, Application US/09082089  
; Patent No. 6100060

; GENERAL INFORMATION:

; APPLICANT: BARNES, MICHAEL

; APPLICANT: TESTA, TANIA

; APPLICANT: KELSELL, DAVID

; TITLE OF INVENTION: No. 6100060e1 Compounds

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RATNER &amp; PRESTIA

; STREET: P.O. BOX 980

; CITY: VALLEY FORGE

; STATE: PA

; COUNTRY: USA

; ZIP: 19482

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/082.089

; FILING DATE: 20-MAY-1998

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: GB 9710737.9

; FILING DATE: 23-MAY-1997

; APPLICATION NUMBER: GB 9803981.1

; FILING DATE: 25-FEB-1998

; APPLICATION NUMBER: GB 9804007.4  
; FILING DATE: 25-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PRESTIA, PAUL F.  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GH-30166  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 363 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-082-089--5

Query Match 50.7%; Score 36; DB 3; Length 363;  
Best Local Similarity 54.5%; Pred. No. 1.9e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 RPK--PQQWF 9  
| | | | |  
Db 152 RPKDLPDRLW 162

RESULT 191  
US-09-082-089-2  
; Sequence 2, Application US/09082089  
; Patent No. 6100060  
; GENERAL INFORMATION:  
; APPLICANT: BARNES, MICHAEL  
; APPLICANT: TESTA, TANIA  
; APPLICANT: KELSELL, DAVID  
; TITLE OF INVENTION: No. 6100060e1 Compounds  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: RATNER & PRESTIA  
; STREET: P.O. BOX 980  
; CITY: VALLEY FORGE  
; STATE: PA  
; COUNTRY: USA  
; ZIP: 19482  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/082,089  
; FILING DATE: 20-MAY-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 9710737.9  
; FILING DATE: 23-MAY-1997  
; APPLICATION NUMBER: GB 9803981.1  
; FILING DATE: 25-FEB-1998  
; APPLICATION NUMBER: GB 9804007.4  
; FILING DATE: 25-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: PRESTIA, PAUL F.  
; REGISTRATION NUMBER: 23,031  
; REFERENCE/DOCKET NUMBER: GH-30166  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 610-407-0700  
; TELEFAX: 610-407-0701  
; TELEX: 846169  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 372 amino acids

; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-09-082-089-2

Query Match 50.7%; Score 36; DB 3; Length 372;  
Best Local Similarity 54.5%; Pred. No. 2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 RPK--PQQWF 9  
| | | | |  
Db 161 RPKDLPDRLW 171

RESULT 192  
US-08-677-049-2  
; Sequence 2, Application US/08677049  
; Patent No. 5858707  
; GENERAL INFORMATION:  
; APPLICANT: Guimaraes, M. Jorge  
; APPLICANT: Bazan, J. Fernando  
; APPLICANT: McCIanahan, Terrill K.  
; APPLICANT: Zlotnik, Albert  
; TITLE OF INVENTION: PURIFIED MAMMALIAN NUCLEOBASE PERMEASES;  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DNAX Research Institute  
; STREET: 901 California Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94304-1104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/677,049  
; FILING DATE: 03-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/000,788  
; FILING DATE: 03-JUL-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ching, Egwin P.  
; REGISTRATION NUMBER: 34,090  
; REFERENCE/DOCKET NUMBER: DX0511  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-852-9196  
; TELEFAX: 415-496-1200  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 611 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-677-049-2

Query Match 50.7%; Score 36; DB 2; Length 611;  
Best Local Similarity 55.6%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PKPQQWF 10  
| | | | |  
Db 296 PSDAPFWL 304

Mon Apr 1 16:34:29 2002

us-09-988-792-2.50pct.ra1

Page 71

Search completed: April 1, 2002, 16:18:46  
Job time: 77 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2002, 16:19:17 ; Search time 23.26 seconds  
(without alignments)  
36.024 Million cell updates/sec

Title: US-09-988-792-2  
Perfect score: 71  
Sequence: 1 RPKPQQWFWM 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

(Total number of hits satisfying chosen parameters: 147

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

(Post-processing: Minimum Match 50%

Maximum Match 100%  
Listing first 1000 summaries

Database :

PIR\_68:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

| Result No. | Score | Query Match | Length | ID       | Description         |
|------------|-------|-------------|--------|----------|---------------------|
| 1          | 48    | 67.6        | 11     | 1 SPHO   | substance P - hors  |
| 2          | 48    | 67.6        | 11     | 1 A60654 | substance P - guin  |
| 3          | 48    | 67.6        | 63     | 2 JC2412 | tachykinin gamma c  |
| 4          | 48    | 67.6        | 72     | 2 I62742 | tachykinin A gamma  |
| 5          | 48    | 67.6        | 72     | 2 JC5455 | preprotachykinin-A  |
| 6          | 48    | 67.6        | 97     | 2 SI2958 | tachykinin delta p  |
| 7          | 48    | 67.6        | 112    | 1 SPRTA  | substance P alpha   |
| 8          | 48    | 67.6        | 115    | 1 SPRBG  | substance P gamma   |
| 9          | 48    | 67.6        | 115    | 2 S47039 | tachykinin 1 precu  |
| 10         | 48    | 67.6        | 129    | 1 SHUB   | neurokinin 1 precu  |
| 11         | 48    | 67.6        | 130    | 1 SPRTB  | substance P beta p  |
| 12         | 48    | 67.6        | 130    | 1 SPBOB  | neurokinin 1 precu  |
| 13         | 48    | 67.6        | 130    | 2 S47038 | tachykinin 1 precu  |
| 14         | 48    | 67.6        | 130    | 2 IS2526 | neurokinin 1 precu  |
| 15         | 45    | 63.4        | 11     | 2 JN0023 | substance P - chic  |
| 16         | 45    | 63.4        | 453    | 2 F81720 | lipid A biosynthes  |
| 17         | 45    | 63.4        | 455    | 2 E71569 | probable acyltrans  |
| 18         | 44    | 62.0        | 437    | 1 B29336 | ubiquinol--cytochr  |
| 19         | 43    | 60.6        | 365    | 2 C81050 | cytochrome c oxida  |
| 20         | 43    | 60.6        | 365    | 2 F81826 | probable cytochrom  |
| 21         | 42    | 59.2        | 378    | 1 A40004 | histidine decarbox  |
| 22         | 41    | 57.7        | 55     | 2 T11538 | H+-transporting AT  |
| 23         | 41    | 57.7        | 158    | 2 H70028 | conserved hypothet  |
| 24         | 41    | 57.7        | 159    | 2 D71500 | hypothetical prote  |
| 25         | 41    | 57.7        | 261    | 2 C83157 | hypothetical prote  |
| 26         | 41    | 57.7        | 286    | 2 S76015 | acetylxlalan estera |
| 27         | 41    | 57.7        | 291    | 1 S76015 | hypothetical prote  |
| 28         | 41    | 57.7        | 304    | 2 A84353 | acetyltransferase   |
| 29         | 41    | 57.7        | 529    | 2 S76167 | hypothetical prote  |

|     |      |      |      |           |                    |
|-----|------|------|------|-----------|--------------------|
| 30  | 40   | 56.3 | 273  | 2 C82182  | hypothetical prote |
| 31  | 40   | 56.3 | 289  | 2 D81085  | HtrB/Msbb family p |
| 32  | 40   | 56.3 | 289  | 2 B81857  | probable acetyltra |
| 33  | 40   | 56.3 | 290  | 2 E71631  | lipid A biosynthes |
| 34  | 40   | 56.3 | 318  | 2 G82350  | lipid A biosynthes |
| 35  | 40   | 56.3 | 343  | 2 T42129  | probable acyltrans |
| 36  | 40   | 56.3 | 363  | 2 S74814  | hypothetical prote |
| 37  | 40   | 56.3 | 665  | 2 S52072  | DmCNGC protein - f |
| 38  | 39   | 55.6 | 548  | 2 S75427  | hypothetical prote |
| 39  | 39   | 54.9 | 55   | 2 T11768  | H+-transporting AT |
| 40  | 39   | 54.9 | 99   | 2 JQ1632  | HCLFI protein - hu |
| 41  | 39   | 54.9 | 208  | 2 S76097  | hypothetical prote |
| 42  | 39   | 54.9 | 341  | 2 B86503  | acyltransferase li |
| 43  | 39   | 54.9 | 361  | 2 A36669  | galactoside 3(4)-L |
| 44  | 39   | 54.9 | 374  | 2 A42270  | alpha (1,3) fucosy |
| 45  | 39   | 54.9 | 412  | 2 D71803  | ubiquinol--cytochr |
| 46  | 39   | 54.9 | 416  | 2 F83010  | probable oxidoredu |
| 47  | 39   | 54.9 | 440  | 1 B29413  | ubiquinol--cytochr |
| 48  | 39   | 54.9 | 452  | 2 F81436  | probable integral  |
| 49  | 39   | 54.9 | 462  | 2 E81551  | lipid A biosynthes |
| 50  | 39   | 54.9 | 467  | 2 B72119  | acyltransferase -  |
| 51  | 39   | 54.9 | 480  | 2 E72682  | hypothetical prote |
| 52  | 39   | 54.9 | 525  | 2 H71365  | probable licc prot |
| 53  | 39   | 54.9 | 1015 | 2 T13062  | CLOCK protein - fr |
| 54  | 39   | 54.9 | 1023 | 2 T13068  | CLOCK protein - fr |
| 55  | 39   | 54.9 | 1027 | 2 T13071  | CLOCK protein - fr |
| 56  | 38   | 53.5 | 105  | 2 D83242  | hypothetical prote |
| 57  | 38   | 53.5 | 323  | 2 A85798  | suppressor of htrB |
| 58  | 38   | 53.5 | 323  | 2 A42608  | (Kdo)2-(lauroyl)-1 |
| 59  | 38   | 53.5 | 328  | 2 A46521  | 52K phosphoprotein |
| 60  | 38   | 53.5 | 330  | 2 A30533  | lymphocyte-specifi |
| 61  | 38   | 53.5 | 330  | 2 I57835  | lymphocyte-specifi |
| 62  | 38   | 53.5 | 363  | 2 JE0111  | lectin-like oxidiz |
| 63  | 38   | 53.5 | 373  | 2 B82697  | rod shape-determin |
| 64  | 38   | 53.5 | 400  | 2 JC4591  | alpha-1,3 fucosylt |
| 65  | 38   | 53.5 | 405  | 2 B36340  | alpha(1,3)-fucosyl |
| 66  | 38   | 53.5 | 412  | 2 C64712  | ubiquinol--cytochr |
| 67  | 38   | 53.5 | 421  | 2 C96806  | unknown protein T5 |
| 68  | 38   | 53.5 | 433  | 2 A57596  | alpha-1,3-fucosylt |
| 69  | 38   | 53.5 | 455  | 2 T04448  | hypothetical prote |
| 70  | 38   | 53.5 | 533  | 2 S62489  | hypothetical prote |
| 71  | 38   | 53.5 | 789  | 2 I59550  | aryl hydrocarbon r |
| 72  | 38   | 53.5 | 901  | 1 WNVNVT  | 104K glycoprotein  |
| 73  | 38   | 53.5 | 909  | 2 S76899  | hypothetical prote |
| 74  | 38   | 53.5 | 1184 | 2 T09484  | cartilage intermed |
| 75  | 38   | 53.5 | 1832 | 2 T31113  | mucin-like glycopr |
| 76  | 37.5 | 52.8 | 298  | 2 D69351  | hypothetical prote |
| 77  | 37.5 | 52.8 | 375  | 2 JQ0846  | DNA-binding protei |
| 78  | 37.5 | 52.8 | 1196 | 1 DNB5V1  | major DNA-binding  |
| 79  | 37.5 | 52.8 | 1196 | 1 DNB5KS  | DNA-binding protei |
| 80  | 37.5 | 52.8 | 1196 | 1 DNB5HF  | DNA-binding protei |
| 81  | 37.5 | 52.8 | 1197 | 1 A48350  | DNA-binding protei |
| 82  | 37.5 | 52.8 | 1204 | 1 DNB5E29 | DNA-binding protei |
| 83  | 37.5 | 52.8 | 1208 | 2 T42574  | DNA-binding protei |
| 84  | 37.5 | 52.8 | 1209 | 1 DNB5C4  | DNA-binding protei |
| 85  | 37   | 52.1 | 11   | 2 S23308  | substance p - rain |
| 86  | 37   | 52.1 | 175  | 2 S50061  | DNA binding protei |
| 87  | 37   | 52.1 | 274  | 2 T39087  | hypothetical prote |
| 88  | 37   | 52.1 | 342  | 2 A54057  | alpha(1,3)-fucosyl |
| 89  | 37   | 52.1 | 365  | 2 S55498  | alpha(1,3/4)-fucos |
| 90  | 37   | 52.1 | 403  | 2 H64861  | hypothetical prote |
| 91  | 37   | 52.1 | 404  | 1 S25953  | ubiquinol--cytochr |
| 92  | 37   | 52.1 | 436  | 2 B82147  | conserved hypothet |
| 93  | 37   | 52.1 | 536  | 2 C82433  | methyl-accepting c |
| 94  | 37   | 52.1 | 655  | 1 A54306  | proprotein convert |
| 95  | 37   | 52.1 | 662  | 1 TOBPUP  | transposase - phag |
| 96  | 37   | 52.1 | 681  | 2 I78558  | hypothetical Brach |
| 97  | 37   | 52.1 | 747  | 2 I39444  | AMP deaminase (EC  |
| 98  | 37   | 52.1 | 1002 | 2 S54252  | deep orange protei |
| 99  | 37   | 52.1 | 1239 | 2 G02750  | DNA-directed DNA p |
| 100 | 37   | 52.1 | 1384 | 2 T26656  | hypothetical prote |
| 101 | 37   | 52.1 | 1635 | 2 T32452  | hypothetical prote |
| 102 | 37   | 52.1 | 2458 | 2 T17420  | probable polyketid |

103 36.5 51.4 537 2 JC4534 cytochrome P450 4F  
104 36.5 51.4 659 2 S77658 hypothetical prote  
105 36.5 51.4 1186 1 DNBBEG DNA-binding protei  
106 36 50.7 11 2 S23306 substance P - Atla  
107 36 50.7 55 2 S45489 H+-transporting AT  
108 36 50.7 103 2 S77270 hypothetical prote  
109 36 50.7 154 2 C75435 hypothetical prote  
110 36 50.7 171 2 E83140 phosphatidylglycer  
111 36 50.7 187 2 S43177 p18 protein - Lels  
112 36 50.7 268 2 T16544 hypothetical prote  
113 36 50.7 280 2 A82185 glycerol-3-phospha  
114 36 50.7 287 2 A75511 conserved hypothet  
115 36 50.7 318 2 I64053 membrane-bound lyt  
116 36 50.7 330 2 S76408 hypothetical prote  
117 36 50.7 347 2 S44995 pectate lyase - Er  
118 36 50.7 353 4 I59347 hypothetical gluta  
119 36 50.7 360 2 S34173 wnt-5c protein - A  
120 36 50.7 365 2 A48914 proto-oncogene Wnt  
121 36 50.7 372 2 S75038 hypothetical prote  
122 36 50.7 379 2 E36470 Wnt-5b protein - m  
123 36 50.7 379 2 D36470 Wnt-5a protein - m  
124 36 50.7 380 2 A71390 ubiquinol--cytochr  
125 36 50.7 380 2 S70394 ubiquinol--cytochr  
126 36 50.7 381 2 S59093 citrate synthase p  
127 36 50.7 473 2 T39028 cytochrome P450 4F  
128 36 50.7 522 2 JC4532 hypothetical prote  
129 36 50.7 605 2 T35047 hypothetical prote  
130 36 50.7 643 2 E69334 acetyl-CoA synthet  
131 36 50.7 656 2 T01573 conserved hypothet  
132 36 50.7 670 2 A75542 earl protein - mal  
133 36 50.7 704 2 H82381 toxin secretion AT  
134 36 50.7 720 2 C85547 probable cytoplasm  
135 36 50.7 774 2 A70010 NADH dehydrogenase  
136 36 50.7 804 2 S61395 probable Na+/H+-ex  
137 36 50.7 804 2 G83814 pled protein - Syn  
138 36 50.7 829 2 S75776 hypothetical prote  
139 36 50.7 861 2 S77086 hypothetical prote  
140 36 50.7 959 2 T14761 hypothetical prote  
141 36 50.7 1038 2 T25033 hypothetical prote  
142 36 50.7 1039 2 T22982 hypothetical prote  
143 36 50.7 1377 2 T19214 UDP-glucose-glyco  
144 36 50.7 1493 2 T16404 hypothetical prote  
145 36 50.7 1674 2 G96736 kakapo gene protei  
146 36 50.7 2396 2 T13714 giant protein p619  
147 35.5 50.0 4861 2 S71752

ALIGNMENTS

RESULT 1  
SPHO substance P - horse  
C:Species: Equus caballus (domestic horse)  
C:Date: 23-Oct-1981 #sequence\_revision 23-Oct-1981 #text\_change 23-Aug-1996  
C:Accession: A01558  
R:Studer, R.O.; Trzeciak, A.; Lergier, W.  
Helv. Chim. Acta 56, 860-866, 1973  
A:Title: Isolierung und Aminosaeuresequenz von Substanz P aus Pferdedarm.  
A:Reference number: A01558  
A:Accession: A01558  
A:Molecule type: protein  
A:Residues: 1-11 <STU>  
C:Superfamily: substance P precursor  
C:Keywords: amidated carboxyl end; hormone  
F:11/Modified site: amidated carboxyl end (Met) #status experimental  
  
Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.096;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RPKPQQFWFLM 11

Db 1 RPKPQQFFGLM 11  
|||||:|  
RESULT 2  
A60654 substance P - guinea pig  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 14-May-1993 #sequence\_revision 27-Jun-1994 #text\_change 08-Dec-1995  
C:Accession: A60654  
R:Murphy, R.  
Neuropeptides 14, 105-110, 1989  
A:Title: Primary amino acid sequence of guinea-pig substance P.  
A:Reference number: A60654; MUID:90044685  
A:Accession: A60654  
A:Molecule type: protein  
A:Residues: 1-11 <WUR>  
C:Superfamily: substance P precursor  
C:Keywords: amidated carboxyl end; neuropeptide; tachykinin  
F:11/Modified site: amidated carboxyl end (Met) #status experimental  
  
Query Match 67.6%; Score 48; DB 1; Length 11;  
Best Local Similarity 81.8%; Pred. No. 0.096;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RPKPQQFWFLM 11  
|||||:|  
Db 1 RPKPQQFFGLM 11  
|||||:|  
RESULT 3  
JC2412 tachykinin gamma chain precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 23-Feb-1995 #sequence\_revision 26-May-1995 #text\_change 17-Mar-1999  
C:Accession: JC2412  
R:Khan, I.; Collins, S.M.  
Biochem. Biophys. Res. Commun. 202, 796-802, 1994  
A:Title: Fourth isoform of preprotachykinin messenger RNA encoding for substance P in  
A:Reference number: JC2411; MUID:94324969  
A:Accession: JC2412  
A:Molecule type: mRNA  
A:Residues: 1-63 <KHA>  
C:Superfamily: substance P precursor  
C:Keywords: amidated carboxyl end  
F:12-21/Product: substance P #status predicted <SUP>  
F:21/Modified site: amidated carboxyl end (Met) (amide in mature form from following  
  
Query Match 67.6%; Score 48; DB 2; Length 63;  
Best Local Similarity 81.8%; Pred. No. 0.54;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
Qy 1 RPKPQQFWFLM 11  
|||||:|  
Db 11 RPKPQQFFGLM 21  
|||||:|  
RESULT 4  
I62742 tachykinin A gamma chain precursor - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
C:Accession: I62742; JC5453  
R:Chiwakata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schulze, W.; Ivell, R.  
Endocrinology 128, 2441-2448, 1991  
A:Title: Tachykinin (substance-P) gene expression in Leydig cells of the human and mo  
A:Reference number: JC5450; MUID:91209287  
A:Accession: I62742  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-72 <RES>

A:Cross-references: GB:M68909; NID:g200469; PIDN:AAA39970.1; PID:g554261  
C:Comment: This protein contains two tachykinin peptide hormone substance-P which is involved in the regulation of the release of substance-P from the nerve terminal.  
C:Genetics:   
A:Gene: gamma-PPT-A  
C:Superfamily: substance P precursor  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-33/Product: substance-P #status predicted <STP>  
F:48-57/Product: neurokinin-A #status predicted <NKA>

Query Match 67.6%; Score 48; DB 2; Length 72;  
Best Local Similarity 81.8%; Pred. No. 0.61;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQQFWFLM 11  
|||||:||  
Db 23 RPKPQQQFFGLM 33

RESULT 5  
JC5455  
substance P alpha precursor - bovine  
N:Alternates: substance P  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-Jul-1997 #sequence\_revision 29-Aug-1997 #text\_change 16-Jul-1999  
C:Accession: JC5455; I45967  
R:Chiwakata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schulze, W.; Ivell, R.  
Endocrinology 128, 2441-2448, 1991  
A:Title: Tachykinin (substance-P) gene expression in Leydig cells of the human and mouse testis.  
A:Reference number: JC5450; MUID:91209287  
A:Accession: JC5455  
A>Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-72 <HT>  
A:Cross-references: GB:M68912; NID:g163593; PIDN:AAA30725.1; PID:g552336  
C:Comment: This protein contains two tachykinin peptide hormone substance-P which is involved in the regulation of the release of substance-P from the nerve terminal.  
C:Genetics:   
A:Gene: PPT-A  
C:Superfamily: substance P precursor  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-33/Product: substance-P #status predicted <STP>  
F:48-57/Product: neurokinin-A #status predicted <NKA>

Query Match 67.6%; Score 48; DB 2; Length 72;  
Best Local Similarity 81.8%; Pred. No. 0.61;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQQFWFLM 11  
|||||:||  
Db 23 RPKPQQQFFGLM 33

RESULT 6  
SI2958  
tachykinin delta precursor - rat  
N:Alternates: neurokinin A; neuropeptide K; gamma-preprotachykinin I precursor; tachykinin  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 18-Feb-1994 #sequence\_revision 10-Nov-1995 #text\_change 16-Jul-1999  
C:Accession: SI2958; JC2413  
R:Harmat, A.J.; Hyde, V.; Chapman, K.  
FEBS Lett. 275, 22-24, 1990  
A:Title: Identification and cDNA sequence of delta-preprotachykinin, a fourth splicing variant of the rat tachykinin gene.  
A:Reference number: SI2958; MUID:91085565  
A:Accession: SI2958  
A:Molecule type: mRNA  
A:Residues: 1-97 <HAR>  
A:Cross-references: GB:X56306; NID:g56067; PIDN:CAA39752.1; PID:g56068  
R:Khan, I.; Collins, S.M.  
Biochem. Biophys. Res. Commun. 202, 796-802, 1994  
A:Title: Fourth isoform of preprotachykinin messenger RNA encoding for substance P in the rat.  
A:Reference number: JC2411; MUID:94324969  
A:Accession: JC2413  
A:Molecule type: mRNA  
A:Residues: 48-92 <KHA>

A:Cross-references: GB:S72369; NID:g632805; PIDN:AAB31499.1; PID:g632806  
C:Superfamily: substance P precursor  
C:Keywords: amidated carboxyl end  
F:59-68/Product: substance P #status predicted <SUP>  
F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following

Query Match 67.6%; Score 48; DB 2; Length 97;  
Best Local Similarity 81.8%; Pred. No. 0.82;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQQFWFLM 11  
|||||:||  
Db 58 RPKPQQQFFGLM 68

RESULT 7  
SPRTA  
substance P alpha precursor - rat  
N:Alternates: preprotachykinin alpha  
N:Contains: substance P  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1988 #sequence\_revision 26-May-1995 #text\_change 18-Jun-1999  
C:Accession: B26590  
R:Krause, J.E.; Chirgwin, J.M.; Carter, M.S.; Xu, Z.S.; Hershey, A.D.  
Proc. Natl. Acad. Sci. U.S.A. 84, 881-885, 1987  
A:Title: Three rat preprotachykinin mRNAs encode the neuropeptides substance P and neuropeptide Y.  
A:Reference number: A94187; MUID:87118268  
A:Accession: B26590  
A:Molecule type: mRNA  
A:Residues: 1-112 <KRA>  
A:Cross-references: GB:M34184; NID:g206329; PIDN:AAA41925.1; PID:g206330  
C:Comment: Alternative splicing of the mRNA for substance P precursor yields the alpha and beta forms.  
C:Superfamily: substance P precursor  
F:1-112/Product: substance P alpha precursor #status predicted <SIG>  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:58-68/Product: substance P #status predicted <SBP>  
F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following

Query Match 67.6%; Score 48; DB 1; Length 112;  
Best Local Similarity 81.8%; Pred. No. 0.95;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQQFWFLM 11  
|||||:||  
Db 58 RPKPQQQFFGLM 68

RESULT 8  
SPRBG  
substance P gamma precursor - rabbit  
N:Alternates: gamma-neuropeptide K; gamma-preprotachykinin I precursor; tachykinin  
N:Contains: neurokinin A; neuropeptide K; substance P  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C:Date: 10-Nov-1992 #sequence\_revision 26-May-1995 #text\_change 18-Jun-1999  
C:Accession: JN0709; A60302; A60200; S18922  
R:Maegert, H.J.; Heitland, A.; Rose, M.; Forssmann, W.G.  
Biochem. Biophys. Res. Commun. 195, 128-131, 1993  
A:Title: Nucleotide sequence of the rabbit gamma-preprotachykinin I cDNA.  
A:Reference number: JN0709; MUID:93371392  
A:Accession: JN0709  
A:Molecule type: mRNA  
A:Residues: 1-115 <MA>  
A:Cross-references: EMBL:X62994; NID:g1565; PIDN:CAA44728.1; PID:g1566  
R:Kage, R.; McGregor, G.P.; Thim, L.; Conlon, J.M.  
Regul. Pept. 18, 346, 1987  
A:Title: gamma-Neuropeptide K: a peptide isolated from rabbit gut that is derived from the pro-gamma-Neuropeptide K.  
A:Reference number: A60302  
A:Accession: A60302  
A:Molecule type: protein

A:Residues: 72-92 <KAG>  
R:Kage, R.; McGregor, G.P.; Thim, L.; Conlon, J.M.  
J. Neurochem. 50, 1412-1417, 1988  
A:Title: Neuropeptide-gamma: a peptide isolated from rabbit intestine that is derived from  
A:Reference number: A60200; MUID:88199570  
A:Accession: A60200  
A:Molecule type: protein  
A:Residues: 72-92 <KA2>  
C:Comment: The gamma alternatively spliced form is processed to yield substance P and neuropeptide K precursor  
C:Superfamily: substance P precursor  
F:1-15/Domain: signal sequence #status predicted <SIG>  
F:58-68/Product: substance P #status predicted <SBP>  
F:72-92/Product: gamma-neuropeptide K #status experimental <NPK>  
F:83-92/Product: neurokinin A #status predicted <NKA>  
F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following glycosylation)  
F:92/Modified site: amidated carboxyl end (Met) (amide in mature form from following glycosylation)

Query Match 67.6% Score 48; DB 1; Length 115;  
Best Local Similarity 81.8%; Pred. No. 0.97;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQWFWM 11  
|||||:|  
Db 58 RPKPQWFGLM 68

RESULT 9  
S47039  
tachykinin 1 precursor - golden hamster  
C:Species: Mesocricetus auratus (golden hamster)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Jul-1999  
C:Accession: S47039  
R:Heitland, A.; Kruhoffer, M.; Juergen Maegert, H.J.; Forssmann, W.G.  
submitted to the EMBL Data Library, July 1994  
A:Reference number: S47038  
A:Accession: S47039  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-115 <HEI>  
A:Cross-references: EMBL:X80663; NID:g520938; PIDN:CAA56692.1; PID:g520939  
C:Superfamily: substance P precursor

Query Match 67.6% Score 48; DB 2; Length 115;  
Best Local Similarity 81.8%; Pred. No. 0.97;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQWFWM 11  
|||||:|  
Db 58 RPKPQWFGLM 68

RESULT 10  
SPHUB  
neurokinin 1 precursor, beta splice form [validated] - human  
N:Alternate names: neurokinin A; neurokinin alpha; neuromedin L; neuropeptide K; preprokinin  
N:Contains: neurokinin 1; neurokinin 1 precursor, alpha splice form; neurokinin 1 precursor  
C:Species: Homo sapiens (man)  
C:Date: 12-Feb-1988 #sequence\_revision 26-May-1995 #text\_change 19-May-2000  
R:Harmar, A.J.; Armstrong, A.; Pascall, J.C.; Chapman, K.; Rosle, R.; Curtis, A.; Goings, F.E.B. Lett. 208, 67-72, 1986  
A:Title: CDNA sequence of human beta-preprotachykinin, the common precursor to substance P and neuropeptide K  
A:Reference number: A24805; MUID:87030957  
A:Accession: A24805  
A:Molecule type: mRNA  
A:Residues: 1-129 <HAR>  
A:Cross-references: GB:M28109; EMBL:X54469; NID:g29482; PIDN:CAA38351.1; PID:g29483  
R:McGregor, G.P.; Conlon, J.M.  
Peptides 11, 907-910, 1990  
A:Title: Characterization of the C-terminal flanking peptide of human beta-preprotachykinin

A:Reference number: A60425; MUID:91133994  
A:Accession: A60425  
A:Molecule type: protein  
A:Residues: 111-126 <MCG>  
R:Theodorsson-Norheim, E.; Joernvall, H.; Andersson, M.; Norheim, I.; Oberg, K.; Jac Eur. J. Biochem. 166, 693-697, 1987  
A:Title: Isolation and characterization of neurokinin A, neurokinin A(3-10) and neurokinin B  
A:Reference number: S00069; MUID:87275962  
A:Accession: S00069  
A:Molecule type: protein  
A:Residues: 98-107 <THE>  
R:Kage, R.; Thim, L.; Creutzfeldt, W.; Conlon, J.M.  
Biochem. J. 253, 203-207, 1988  
A:Title: Post-translational processing of preprotachykinins. Isolation of protachykinin  
A:Reference number: S03033; MUID:88339887  
A:Accession: S03033  
A:Molecule type: protein  
A:Residues: 20-30 <KAG>  
R:Chiwakata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schulze, W.; Ivell, R.  
Endocrinology 128, 2441-2448, 1991  
A:Title: tachykinin (substance-P) gene expression in Leydig cells of the human and mouse  
A:Reference number: JC5450; MUID:91209287  
A:Accession: JC5451  
A:Status: translation not shown; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 36-73, 89-122 <CHII>  
A:Cross-references: GB:M68907; NID:g190292; PIDN:AAA60160.1; PID:g553619  
A:Accession: JC5450  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 36-86, 'P', 88-122 <CHII>  
A:Cross-references: GB:M68906; NID:g190290; PIDN:AAA60159.1; PID:g553618  
R:Tan, A.; Too, H.P.  
submitted to GenBank, October 1995  
A:Reference number: A59269  
A:Accession: A59269  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-129 <TAN>  
A:Cross-references: GB:037529; NID:g1017792; PIDN:AA79195.1; PID:g1017793  
A:Experimental source: tissue brain cortex  
R:Lai, J.P.; Douglas, S.D.; Rappaport, E.; Wu, J.M.; Ho, W.Z.  
submitted to GenBank, February 1998  
A:Description: Identification of a delta isoform of preprotachykinin mRNA in human  
A:Reference number: A59270  
A:Accession: A59270  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 36-96, 'M', 116-118 <LAI1>  
A:Cross-references: GB:AF050656; NID:g3098594; PIDN:AAC15702.1; PID:g3098595  
A:Experimental source: alpha splice form; tissue blood; tissue brain; cell type monoclonal  
C:Comment: This protein is processed to produce the tachykinin peptide hormones neurokinin K.  
C:Genetics:  
A:Gene: GDB:TAC2; NKNA; PPT-A  
A:Cross-references: GDB:119452; OMIM:162320  
A:Map position: 7q21-q22  
C:Superfamily: substance P precursor  
C:Keywords: alternative splicing; amidated carboxyl end; hormone; neuropeptide; tachykinin  
F:1-129/Product: neurokinin 1 precursor, beta splice form #status predicted <SPB>  
F:1-96, 'M', 116-118/Product: neurokinin 1 precursor, alpha splice form #status predicted <S>  
F:1-73, 89-129/Product: neurokinin 1 precursor, gamma splice form #status predicted <S>  
F:1-73, 89-96, 'M', 116-122/Product: neurokinin 1 precursor, alpha splice form #status predicted <S>  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-57/Domain: amino-terminal propeptide #status predicted <PRO>  
F:58-68/Product: neurokinin 1 #status experimental <NK1>



F:72-107/Product: neuropeptide K #status predicted <NEK>  
F:98-107/Product: neurokinin 2 #status experimental <NK2>  
F:100-107/Product: neurokinin 2(3-10) #status experimental <NK23>  
F:101-107/Product: neurokinin 2(4-10) #status experimental <NK24>  
F:111-126/Domain: carboxyl-terminal propeptide #status experimental <CTP>  
F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following gly  
F:107/Modified site: amidated carboxyl end (Met) (amide in mature form from following gly

Query Match 67.6%; Score 48; DB 1; Length 129;  
Best Local Similarity 81.8%; Pred. NO. 1.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RPKPOQWFWM 11  
|||||:|  
DB 58 RPKPOQFGLM 68

RESULT 11  
SPRTB  
Substance P beta precursor - rat  
N:Alternate names: preprotachykinin beta; preprotachykinin gamma; substance K  
N:Contains: neurokinin A; substance P; substance P gamma precursor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-1988 #sequence\_revision 26-May-1995 #text\_change 18-Jun-1999  
C:Accession: A37163; A26590; C26590; A25067; JC2411  
R:Carter, M.S.; Krause, J.E.  
J. Neurosci. 10, 2203-2214, 1990  
A:Title: Structure, expression, and some regulatory mechanisms of the rat preprotachykin  
A:Reference number: A37163; MUID:90331040  
A:Accession: A37163  
A:Molecule type: DNA  
A:Residues: 1-130 <CAR>  
A:Cross-references: GB:M34159; GB:M34160; GB:M34162; NID:g206334; PIDN:AAA41926.1; PID:9  
R:Krause, J.E.; Chirgwin, J.M.; Carter, M.S.; Xu, Z.S.; Hershey, A.D.  
Proc. Natl. Acad. Sci. U.S.A. 84, 881-885, 1987  
A:Title: Three rat preprotachykinin mRNAs encode the neuropeptides substance P and neuro  
A:Reference number: A94187; MUID:87118268  
A:Accession: A26590  
A:Molecule type: mRNA  
A:Residues: 1-130 <KRA>  
A:Cross-references: GB:M15191; NID:g206341; PIDN:AAA41928.1; PID:g206342; GB:M35277  
A:Accession: C26590  
A:Molecule type: mRNA  
A:Residues: 1-73, 89-130 <KR2>  
A:Cross-references: GB:M34183; NID:g206343; PIDN:AAA41929.1; PID:g206344  
R:Kawaguchi, Y.; Hoshimaru, M.; Nawa, H.; Nakanishi, S.  
Biochem. Biophys. Res. Commun. 139, 1040-1046, 1986  
A:Title: Sequence analysis of cloned cDNA for rat substance P precursor: existence of a  
A:Reference number: A25067; MUID:87025808  
A:Accession: A25067  
A:Molecule type: mRNA  
A:Residues: 1-73, 89-130 <KAW>  
A:Cross-references: GB:M14312; NID:g206339; PIDN:AAA41927.1; PID:g206340  
R:Khan, I.; Collins, S.M.  
Biochem. Biophys. Res. Commun. 202, 796-802, 1994  
A:Title: Fourth isoform of preprotachykinin messenger RNA encoding for substance P in th  
A:Reference number: JC2411; MUID:94324969  
A:Accession: JC2411  
A:Molecule type: mRNA  
A:Residues: 48-110 <KHA>  
A:Experimental source: intestine  
A:Comment: Alternative splicing of the mRNA for substance P precursor yields the beta an  
C:Comment: The beta and gamma forms are processed to yield substance P and neurokinin A  
C:Genetics:  
A:Introns: 41/3; 74/1; 89/1; 97/1; 115/1  
C:Superfamily: substance P precursor  
C:Keywords: alternative splicing; amidated carboxyl end; hormone; neuropeptide; tachykin  
F:1-130/Product: substance P beta precursor #status predicted <PREB>  
F:1-73, 89-130/Domain: signal sequence #status predicted <SIG>  
F:58-68/Product: substance P #status predicted <SBP>  
F:98-107/Product: neurokinin A #status predicted <NKA>

F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following  
F:107/Modified site: amidated carboxyl end (Met) (amide in mature form from following

Query Match 67.6%; Score 48; DB 1; Length 130;  
Best Local Similarity 81.8%; Pred. NO. 1.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 RPKPOQWFWM 11  
|||||:|  
DB 58 RPKPOQFGLM 68

RESULT 12  
SPROB  
neurokinin 1 precursor, beta splice form [validated] - bovine  
N:Alternate names: neurokinin A; preprotachykinin; substance K; substance P  
N:Contains: neurokinin 1; neurokinin 1 precursor, alpha splice form; neurokinin 1 pre  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Feb-1984 #sequence\_revision 19-Feb-1984 #text\_change 16-Jun-2000  
C:Accession: A05093; A01559; A01557; B25067; A61460; JC5454; I45966  
R:Nawa, H.; Kotani, H.; Nakanishi, S.  
Nature 312, 729-734, 1984  
A:Title: Tissue-specific generation of two preprotachykinin mRNAs from one gene by al  
A:Reference number: A05093; MUID:85086245  
A:Accession: A05093  
A:Molecule type: DNA  
A:Residues: 1-130 <NAWL>  
A:Cross-references: GB:X02351; GB:M14786; NID:g655; PIDN:CAA24942.1; PID:g1197197  
R:Nawa, H.; Hirose, T.; Takashima, H.; Inayama, S.; Nakanishi, S.  
Nature 306, 32-36, 1983  
A:Title: Nucleotide sequences of cloned cDNAs for two types of bovine brain substance  
A:Reference number: A93318; MUID:84039802  
A:Accession: A01559  
A:Molecule type: mRNA  
A:Residues: 1-130 <NAW2>  
A:Cross-references: GB:X00075; NID:g758; PIDN:CAA24939.1; PID:g759  
A:Accession: A01557  
A:Molecule type: mRNA  
A:Residues: 1-96, 'M', 116-130 <NAW3>  
A:Cross-references: GB:X00076; NID:g762; PIDN:CAA24942.1; PID:g763  
R:Kawaguchi, Y.; Hoshimaru, M.; Nawa, H.; Nakanishi, S.  
Biochem. Biophys. Res. Commun. 139, 1040-1046, 1986  
A:Title: Sequence analysis of cloned cDNA for rat substance P precursor: existence of  
A:Reference number: A25067; MUID:87025808  
A:Accession: B25067  
A:Molecule type: mRNA  
A:Residues: 1-73, 89-130 <KAW>  
R:McGregor, G.P.; Kage, R.; Thim, L.; Conlon, J.M.  
J. Neurochem. 53, 1871-1877, 1989  
A:Title: Quantitation and characterization of peptides from the C-terminal flanking r  
A:Reference number: A61460; MUID:90039314  
A:Accession: A61460  
A:Molecule type: protein  
A:Residues: 111-126 <MCG>  
A:Experimental source: corpus striatum  
R:Chiwakata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schulze, W.; Ivell, R.  
Endocrinology 128, 2441-2448, 1991  
A:Title: Tachykinin (substance-P) gene expression in Leydig cells of the human and mo  
A:Reference number: JC5450; MUID:91209287  
A:Accession: JC5454  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 36-120, 'A', 122 <CHI>  
A:Cross-references: GB:M68911; NID:g163591; PIDN:AAA30724.1; PID:g552335  
C:Comment: The protein is processed to produce neurokinin 1 (substance P) and neuroki  
C:Genetics:  
A:Gene: PPT-A  
A:Introns: 41/3; 74/1; 89/1; 97/1; 115/1  
C:Superfamily: substance P precursor  
C:Keywords: alternative splicing; amidated carboxyl end; hormone; neuropeptide; tachy  
F:1-130/Product: neurokinin 1 precursor, beta splice form #status predicted <SPB>  
F:1-96, 'M', 116-130/Product: neurokinin 1 precursor, alpha splice form #status predict

F:1-73-89-130/Product: neurokinin 1 precursor, gamma splice form #status predicted <SPG>  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-57/Domain: amino-terminal propeptide #status predicted <PRO>  
F:58-68/Product: neurokinin 1 #status experimental <SBP>  
F:98-107/Product: neurokinin 2 #status predicted <NEK>  
F:111-126/Domain: carboxyl-terminal propeptide #status experimental <CTP>  
F:111-126/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following gly  
F:107/Modified site: amidated carboxyl end (Met) (amide in mature form from following gly

Query Match 67.6%; Score 48; DB 1; Length 130;  
Best Local Similarity 81.8%; Pred. No. 1.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:||  
Db 58 RPKPQQFFGLM 68

RESULT 13  
S47038  
tachykinin 1 precursor - golden hamster  
C:Species: Mesocricetus auratus (golden hamster)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 16-Jul-1999  
C:Accession: S47038  
R:Heitland, A.; Kruhoffer, M.; Juergen Maegert, H.J.; Forssmann, W.G.  
submitted to the EMBL Data Library, July 1994  
A:Reference number: S47038  
A:Accession: S47038  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-130 <HEI>  
A:Cross-references: EMBL:X80662; NID:g520917; PIDN:CAA56691.1; PID:g520918  
C:Superfamily: substance P precursor

Query Match 67.6%; Score 48; DB 2; Length 130;  
Best Local Similarity 81.8%; Pred. No. 1.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:||  
Db 58 RPKPQQFFGLM 68

RESULT 14  
I52526  
neurokinin 1 precursor - mouse  
N:Alternate names: neurokinin A; preprotachykinin; substance K; substance P  
N:Contains: neurokinin 1; neurokinin 2  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 26-May-2000  
C:Accession: I52526; JC5452; I62741  
R:Kako, K.; Munekata, E.; Hosaka, M.; Murakami, K.; Nakayama, K.  
Biomed. Res. 14, 253-259, 1993  
A:Title: Cloning and sequence analysis of mouse cDNAs encoding preprotachykinin A and B.  
A:Reference number: I52526  
A:Accession: I52526  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-130 <KAK>  
A:Cross-references: GB:D17584; NID:g407345; PIDN:BAA04508.1; PID:g435121  
R:Chiwakata, C.; Brackmann, B.; Hunt, N.; Davidoff, M.; Schulze, W.; Ivell, R.  
Endocrinology 128, 2441-2448, 1991  
A:Title: Tachykinin (substance-P) gene expression in Leydig cells of the human and mouse  
A:Reference number: JC5450; MUID:91209287  
A:Accession: JC5452  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 36-122 <CHI>  
A:Cross-references: GB:M68908; NID:g200467; PIDN:AAA39969.1; PID:g554260  
C:Genetics:  
A:Gene: PPT-A

C:Superfamily: substance P precursor  
C:Keywords: amidated carboxyl end  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-57/Domain: amino-terminal propeptide #status predicted <PRO>  
F:58-68/Product: neurokinin 1 #status predicted <NK1>  
F:98-107/Product: neurokinin 2 #status predicted <NK2>  
F:111-126/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
F:111-126/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
F:68/Modified site: amidated carboxyl end (Met) (amide in mature form from following  
F:107/Modified site: amidated carboxyl end (Met) (amide in mature form from following

Query Match 67.6%; Score 48; DB 2; Length 130;  
Best Local Similarity 81.8%; Pred. No. 1.1;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||||:||  
Db 58 RPKPQQFFGLM 68

RESULT 15  
JN0023  
substance P - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 07-Sep-1990 #sequence\_revision 07-Sep-1990 #text\_change 11-Jul-1997  
C:Accession: JN0023  
R:Conlon, J.M.; Katsoulis, S.; Schmidt, W.E.; Thim, L.  
Regul. Pept. 20, 171-180, 1988  
A:Title: [Arg3]substance P and neurokinin A from chicken small intestine.  
A:Reference number: JN0023; MUID:88204263  
A:Accession: JN0023  
A:Molecule type: protein  
A:Residues: 1-11 <CON>  
C:Superfamily: substance P precursor  
C:Keywords: amidated carboxyl end; tachykinin  
F:11/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 63.4%; Score 45; DB 2; Length 11;  
Best Local Similarity 72.7%; Pred. No. 0.28;  
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|||:|:|  
Db 1 RRPQPQFFGLM 11

RESULT 16  
F81720  
lipid A biosynthesis lauroyl acyltransferase, probable TC0278 [imported] - Chlamydia  
C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: F81720  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
A:Reference number: A81500; MUID:20150255  
A:Accession: F81720  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-453 <TET>  
A:Cross-references: GB:AE002295; GB:AE002160; NID:g7190314; FIDN:AAF39146.1; PID:g719  
A:Experimental source: strain Nigg (MoPn)  
C:Genetics:  
A:Gene: TC0278

Query Match 63.4%; Score 45; DB 2; Length 453;  
Best Local Similarity 75.0%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPQQWFWM 10

Db 303 KPEQWLWL 310  
||:|:| ||  
RESULT 17  
E71569  
probable acyltransferase - Chlamydia trachomatis (serotype D, strain UW3/Cx)  
C:Species: Chlamydia trachomatis  
C:Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 08-Oct-1999  
C:Accession: E71569  
R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell, Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis  
A:Reference number: A71570; MUID:99000809  
A:Accession: E71569  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-455 <ARN>  
A:CROSS-references: GB:AE001275; GB:AE001273; NID:g3328388; PIDN:AAC67600.1; PID:g332839  
A:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: htrB  
Query Match 63.4%; Score 45; DB 2; Length 455;  
Best Local Similarity 75.0%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 3 KPEQWFWL 10  
||:|:| ||  
Db 303 KPEQWLWL 310  
RESULT 18  
B29336  
ubiquinol--cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Rhodobacter capsulatus  
C:Species: Rhodobacter capsulatus  
C:Date: 31-Dec-1988 #sequence\_revision 22-Jul-1994 #text\_change 03-Mar-2000  
C:Accession: B29336; B25405; S09373  
R:Davidson, E.; Daidal, F.  
J. Mol. Biol. 195, 13-24, 1987  
A:Title: Primary structure of the bc-1 complex of Rhodospseudomonas capsulata. Nucleotide sequence and transcription of the fbc operon from Rhodospseudomonas sphaeroides  
A:Reference number: A92938; MUID:88011223  
A:Accession: B29336  
A:Molecule type: DNA  
A:Residues: 1-437 <DAV>  
A:CROSS-references: EMBL:X05630; NID:g46093; PIDN:CAA29117.1; PID:g46095  
R:Gabellini, N.; Seibald, W.  
Eur. J. Biochem. 154, 569-579, 1986  
A:Title: Nucleotide sequence and transcription of the fbc operon from Rhodospseudomonas sphaeroides  
A:Reference number: A91162; MUID:86136096  
A:Note: source is designated as Rhodospseudomonas sphaeroides  
A:Accession: B25405  
A:Molecule type: DNA  
A:Residues: 1-66, 'ID', 69-280, 'I', 282-437 <GAB>  
A:CROSS-references: EMBL:X03476; NID:g46007; PIDN:CAA27195.1; PID:g46009  
C:Genetics:  
A:Gene: fbcB; petB  
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol  
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; oxidoreductase  
F:26-381/Domain: cytochrome b homology <CBH>  
F:26-225/Domain: cytochrome b6 homology <CB6>  
F:51-67/Domain: transmembrane #status predicted <TM1>  
F:96-114/Domain: transmembrane #status predicted <TM2>  
F:134-150/Domain: transmembrane #status predicted <TM3>  
F:146-193/Domain: periplasmic #status predicted <PER1>  
F:195-217/Domain: transmembrane #status predicted <TM4>  
F:245-381/Domain: plastoquinol--plastocyanin reductase 17K protein homology <17K>  
F:253-269/Domain: transmembrane #status predicted <TM5>  
F:270-329/Domain: periplasmic #status predicted <PER2>  
F:330-346/Domain: transmembrane #status predicted <TM6>  
F:365-383/Domain: transmembrane #status predicted <TM7>

F:395-411/Domain: transmembrane #status predicted <TM8>  
F:97,198/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted  
F:111,212/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted  
Query Match 62.0%; Score 44; DB 1; Length 437;  
Best Local Similarity 54.5%; Pred. No. 15;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 RPKQOWFWLM 11  
||:|:| ||  
Db 360 RPKFRWFWFL 370  
RESULT 19  
C81050  
cytochrome c oxidase, chain III NMB1723 [imported] - Neisseria meningitidis (strain M  
C:Species: Neisseria meningitidis  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: C81050  
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B. ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M. Science 287, 1809-1815, 2000  
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.  
A:Reference number: A81000; MUID:20175755  
A:Accession: C81050  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-365 <TET>  
A:CROSS-references: GB:AE002522; GB:AE002098; NID:g7226972; PIDN:AAF42068.1; PID:g722  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB1723  
Query Match 60.6%; Score 43; DB 2; Length 365;  
Best Local Similarity 56.7%; Pred. No. 18;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 2 RPKQOWFWL 10  
||:|:| ||  
Db 62 PLPRWFWL 70  
RESULT 20  
F81826  
probable cytochrome c NMA1977 [imported] - Neisseria meningitidis (strain Z2491 serog  
C:Species: Neisseria meningitidis  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: F81826  
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; MO  
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre Nature 404, 502-506, 2000  
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491  
A:Reference number: A81775; MUID:20222556  
A:Accession: F81826  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-365 <PAR>  
A:CROSS-references: GB:AL162757; GB:AL157959; NID:g7380371; PIDN:CAB85197.1; PID:g738  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: NMA1976; NMA1977  
Query Match 60.6%; Score 43; DB 2; Length 365;  
Best Local Similarity 56.7%; Pred. No. 18;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
Qy 2 RPKQOWFWL 10  
||:|:| ||

Db 62 PLPRWFNL 70

RESULT 21

A40004 histidine decarboxylase (EC 4.1.1.22) - Enterobacter aerogenes

C:Species: Enterobacter aerogenes

C>Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 18-Jun-1999

C:Accession: A40004

R:Kamath, A.V.; Vaaler, G.L.; Snell, E.E.

J. Biol. Chem. 266, 9432-9437, 1991

A:Title: Pyridoxal phosphate-dependent histidine decarboxylases. Cloning, sequencing, and enzymes.

A:Reference number: A40004; MUID:91236707

A:Accession: A40004

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-378 <RAM>

A:Cross-references: GB:M62745; NID:g435593; PIDN:AAA24802.1; PID:g435594

C:Superfamily: Klebsiella histidine decarboxylase

C:Keywords: carbon-carbon lyase; carboxy-lyase; phosphoprotein; pyridoxal phosphate

F:233/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 59.2%; Score 42; DB 1; Length 378;

Best Local Similarity 62.5%; Pred. No. 26;

Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQWFNL 9

||| :| |

Db 329 PKPSEWNL 336

RESULT 22

T11538

H+-transporting ATP synthase (EC 3.6.1.34) protein 8 - spiny dogfish mitochondrion

C:Species: mitochondrion Squalus acanthias (spiny dogfish)

C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C:Accession: T11538

R:Rasmussen, A.S.; Arnason, U.

J. Mol. Evol. 48, 118-123, 1999

A:Title: Phylogenetic studies of complete mitochondrial DNA molecules place cartilaginous

A:Reference number: Z17281; MUID:99091711

A:Accession: T11538

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-55 <RAS>

A:Cross-references: EMBL:Y18134; NID:g4186095; PIDN:CAA77053.1; PID:g4186100

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SGC1

C:Superfamily: H+-transporting ATP synthase protein 8

C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion; ox

Query Match 57.7%; Score 41; DB 2; Length 55;

Best Local Similarity 55.6%; Pred. No. 5.6;

Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQWFNL 9

:||| :| |

Db 44 KKPPEPWNW 52

RESULT 23

H70028

conserved hypothetical protein yvaW - Bacillus subtilis

C:Species: Bacillus subtilis

C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999

C:Accession: H70028

R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berta

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, K.; Tosa, V.; Uchiya  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Zanchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis

A:Reference number: A69580; MUID:98044033

A:Accession: H70028

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-158 <KUN>

A:Cross-references: GB:Z99121; GB:AL009126; NID:g2635827; PIDN:CAB15380.1; PID:ell860

A:Experimental source: strain 168

C:Genetics:

A:Gene: yvaW

Query Match 57.7%; Score 41; DB 2; Length 158;

Best Local Similarity 50.0%; Pred. No. 16;

Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQWFNL 10

||| :| |

Db 131 RQKPLSWYI 140

RESULT 24

D71500

hypothetical protein CT556 - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C>Date: 13-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 18-Aug-2000

C:Accession: D71500

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, F.; Aravind, L.; Mitche

Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t

A:Reference number: A71570; MUID:99000809

A:Accession: D71500

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-159 <ARN>

A:Cross-references: GB:AE001326; GB:AE001273; NID:g3328990; PILEN:AA68158.1; PID:g332

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: CT556

C:Superfamily: conserved hypothetical protein TC0844

Query Match 57.7%; Score 41; DB 2; Length 159;

Best Local Similarity 66.7%; Pred. No. 16;

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQWFNL 10

||| :| |

Db 89 PAFSQWDNL 97

RESULT 25

C83157

hypothetical protein PA3907 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: C83157

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warriner, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337  
A:Accession: C83157  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-261 <STO>  
A:Cross-references: GB:AF004808; GB:AF004091; NID:9950086; PIDN:AAG07294.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA3907

Query Match 57.7%; Score 41; DB 2; Length 261;  
Best Local Similarity 62.5%; Pred. No. 26;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PKPQQWF 9  
| | : |||  
Db 126 PPHFW 133

RESULT 26  
B82564  
acetylxllyan esterase XF2395 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: For a complete list of authors see reference number A59328 below

A:Accession: B82564  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-286 <SIM>  
A:Cross-references: GB:AE004048; GB:AE003849; NID:g9107566; PIDN:AAF85194.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briñones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tshukako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Contents: annotation  
C:Genetics:  
A:Gene: XF2395

Query Match 57.7%; Score 41; DB 2; Length 286;  
Best Local Similarity 62.5%; Pred. No. 29;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQQWF 11  
| | | | :  
Db 165 PQHFW 172

RESULT 27  
S76015  
hypothetical protein - Synecocystis sp. (strain PCC 6803)

C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000

C:Accession: S76015  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;  
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocys  
s.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76015  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-291 <KAN>

A:Cross-references: EMBL:D64006; GB:AB001339; NID:g1001291; PID:g1001372  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: Aquifex aeolicus cysQ protein

Query Match 57.7%; Score 41; DB 1; Length 291;  
Best Local Similarity 50.0%; Pred. No. 29;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQWF 11  
| | | | :  
Db 88 PLPQDW 97

RESULT 28  
A84353  
acetyltransferase homolog [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: A84353  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky  
J.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483

A:Accession: A84353  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-304 <STO>  
A:Cross-references: GB:AE004437; NID:g10581454; PIDN:AAG20189.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: Yyai

Query Match 57.7%; Score 41; DB 2; Length 304;  
Best Local Similarity 66.7%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
| | | | |  
Db 96 RSKPLQWL 104

RESULT 29  
S76167  
hypothetical protein - Synecocystis sp. (strain PCC 6803)

C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76167  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocys  
s.

A:Reference number: S74322; MUID:97061201  
A:Accession: S76167  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-529 <KAN>

A:Cross-references: EMBL:D90914; GB:AB001339; NID:g1653477; PIDN:BAAL8426.1; PID:d101  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

```
Query Match          57.7%; Score 41; DB 2; Length 529;
Best Local Similarity 70.0%; Pred. No. 53;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 10
   |||||
Db 368 RPLPQDWF 377

RESULT 30
C82182
hypothetical protein VC1577 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: C82182
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
  Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers,
  L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: C82182
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-273 <HEI>
A:Cross-references: GB:AE004235; GB:AE003852; NID:g9656082; PIDN:AAF94731.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC1577
A:Map position: 1

Query Match          56.3%; Score 40; DB 2; Length 273;
Best Local Similarity 71.4%; Pred. No. 39;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQWF 10
   |||
Db 262 PEQIWL 268

RESULT 31
D81085
HtrB/MsbB family protein NMB1418 [imported] - Neisseria meningitidis (strain MC58 serog
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C:Accession: D81085
R:Rettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
  Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
  ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
  A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: A81000; MUID:20175755
A:Accession: D81085
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <TET>
A:Cross-references: GB:AE002491; GB:AE002098; NID:g7226655; PIDN:AAF41779.1; PID:g722665
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1418

Query Match          56.3%; Score 40; DB 2; Length 289;
Best Local Similarity 60.0%; Pred. No. 41;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 10
   |||||
Db 265 REHPEQI 274

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: G82350
```

```
RESULT 32
B81857
probable acetyltransferase NMA1630 [imported] - Neisseria meningitidis (strain 22491
C:Species: Neisseria meningitidis
C:Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C:Accession: B81857
R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
  Holroyd, S.; Jagels, K.; Leather, S.; Mungall, K.; Quail, M.A.; Rajandre
  Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis 22491
A:Reference number: A81775; MUID:20222556
A:Accession: B81857
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-289 <PAR>
A:Cross-references: GB:AL162756; GB:AL157959; NID:g7380091; PIDN:CA884858.1; PID:g738
A:Experimental source: serogroup A, strain 22491
C:Genetics:
A:Gene: NMA1630

Query Match          56.3%; Score 40; DB 2; Length 289;
Best Local Similarity 60.0%; Pred. No. 41;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 10
   |||||
Db 265 REHPEQI 274

RESULT 33
E71631
Lipid A biosynthesis lauroyl acyltransferase (htrB) RP718 - Rickettsia prowazekii
C:Species: Rickettsia prowazekii
C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C:Accession: E71631
R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark
  Nature 396, 133-140, 1998
A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A:Reference number: A71630; MUID:99039499
A:Accession: E71631
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-290 <AND>
A:Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15149.1; PID:e134
A:Experimental source: strain Madrid E
C:Genetics:
A:Gene: htrB; RP718

Query Match          56.3%; Score 40; DB 2; Length 290;
Best Local Similarity 55.8%; Pred. No. 41;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9
   |||||
Db 275 KQNPQWF 283

RESULT 34
G82350
Lipid A biosynthesis lauroyl acyltransferase VC0213 [imported] - Vibrio cholerae (str
C:Species: Vibrio cholerae
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: G82350
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
  Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
  L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833
A:Accession: G82350
```

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-318 <HEI>  
A:Cross-references: GB:AB004111; GB:AE003852; NID:g9654614; PIDN:AAF93389.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype E1 Tor  
C:Genetics:  
A:Gene: VC0213  
A:Map position: 1

Query Match 56.3%; Score 40; DB 2; Length 318;  
Best Local Similarity 71.4%; Pred. No. 45;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQWFWL 10  
|:|:| |  
DB 294 PQQWWWL 300

RESULT 35  
T42129  
C:Species: Escherichia coli  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Nov-2000  
C:Accession: T42129; T00321  
R:Burland, V.; Shao, Y.; Perna, N.T.; Plunkett, G.; Sofia, H.J.; Blattner, F.R.  
Nucleic Acids Res. 26, 4196-4204, 1998  
A:Title: The complete DNA sequence and analysis of the large virulence plasmid of Escherichia coli O157:H7  
A:Reference number: Z22068; MUID:98391744  
A:Accession: T42129  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-343 <BUR>  
A:Cross-references: EMBL:AF074613; PIDN:AAC70097.1  
A:Experimental source: strain EDL933; serotype O157:H7  
R:Makino, K.; Ishii, K.; Yasunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota, S.; Shinagawa, H.  
DNA Res. 5, 1-9, 1998  
A:Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic E. coli O157:H7  
A:Reference number: Z14127; MUID:98290540  
A:Accession: T00321  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 12-343 <MAK>  
A:Cross-references: EMBL:AB011549; NID:g4589740; PIDN:BAA31840.1; PID:g3337081  
A:Experimental source: strain EHEC O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Genome: plasmid pO157  
A:Note: L7029  
C:Keywords: acyltransferase

Query Match 56.3%; Score 40; DB 2; Length 343;  
Best Local Similarity 45.5%; Pred. No. 49;  
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
|:|:| |:  
DB 319 RPHPQYTWIL 329

RESULT 36  
S74814  
hypothetical protein slr1737 - Synecocystis sp. (strain PCC 6803)  
C:Species: Synecocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis sp.

A:Reference number: S74322; MUID:97061201  
A:Accession: S74814  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-363 <KAN>  
A:Cross-references: EMBL:D90909; GB:AB001339; NID:g1652844; PIDN:BAAL7775.1; PID:dl01  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 56.3%; Score 40; DB 2; Length 363;  
Best Local Similarity 71.4%; Pred. No. 52;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQWFWL 10  
|:|:| |  
DB 211 PSRWFWL 217

RESULT 37  
S52072  
DMCNC protein - fruit fly (Drosophila sp.)  
C:Species: Drosophila sp.  
C:Date: 14-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 16-Jul-1999  
C:Accession: S52072  
R:Baumann, A.; Frings, S.; Godde, M.; Seifert, R.; Kaupp, U.B.  
EMBO J. 13, 5040-5050, 1994  
A:Title: Primary structure and functional expression of a Drosophila cyclic nucleotidic acid phosphodiesterase  
A:Reference number: S52072; MUID:95045396  
A:Accession: S52072  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-665 <BAU>  
C:Genetics:  
A:Gene: FlyBase:Cng  
A:Cross-references: FlyBase:FBgn0014462  
C:Superfamily: cyclic nucleotide-gated channel; cAMP receptor protein cyclic nucleotid  
F.429-553/Domain: cAMP receptor protein cyclic nucleotide-binding domain homology <CA

Query Match 56.3%; Score 40; DB 2; Length 665;  
Best Local Similarity 75.0%; Pred. No. 94;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 8  
|:|:| |  
DB 52 RPKPPDWF 59

RESULT 38  
B75427  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: B75427  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: B75427  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-548 <WHI>  
A:Cross-references: GB:AE001967; GB:AE005513; NID:g6458915; PIDN:AAF10758.1; PID:g645  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1179  
A:Map position: 1

Query Match 55.6%; Score 39.5; DB 2; Length 548;  
Best Local Similarity 70.0%; Pred. No. 93;

Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
QY 1 RPKPQOWFWL 10  
||| |||  
Db 521 RPVPQEW-WL 529  
RESULT 39  
T11768  
H+-transporting ATP synthase (EC 3.6.1.34) protein 8 - Mustelus manazo mitochondrion  
C:Species: mitochondrion Mustelus manazo  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 20-Jun-2000  
C:Accession: T11768  
R:Cao, Y.; Wadell, P.J.; Okada, N.; Hasegawa, M.  
Mol. Biol. Evol. 15, 1637-1646, 1998  
A:Title: The complete mitochondrial DNA sequence of the shark (Mustelus manazo): Evaluation of the complete mitochondrial DNA sequence  
A:Reference number: Z17338; MUID:99083431  
A:Accession: T11768  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-55 <CAO>  
A:Cross-references: EMBL:AB015962; PIDN:BAA33040.1  
A:Experimental source: liver  
C:Genetics:  
A:Genome: mitochondrion  
A:Note: Atp8  
C:Superfamily: H+-transporting ATP synthase protein 8  
C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex; mitochondrion; oxid

Query Match 54.9%; Score 39; DB 2; Length 55;  
Best Local Similarity 55.6%; Pred. No. 11;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 RPKPQOWFW 9  
||| |||  
Db 44 KPKPNPNW 52  
RESULT 40  
JQ1632  
HCLF1 protein - human herpesvirus 6 (strain U1102)  
C:Species: human herpesvirus 6  
C>Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 08-Oct-1999  
C:Accession: JQ1632  
R:Thomson, B.J.; Honess, R.W.  
J. Gen. Virol. 73, 1649-1660, 1992  
A:Title: The right end of the unique region of the genome of human herpesvirus 6 U1102  
A:Reference number: PQ0406; MUID:92333248  
A:Accession: JQ1632  
A:Molecule-type: DNA  
A:Residues: 1-99 <THO>  
A:Cross-references: DBJ:D11134; NID:9221448; PIDN:BAA01908.1; PID:d1002385; PID:g221454

Query Match 54.9%; Score 39; DB 2; Length 99;  
Best Local Similarity 50.0%; Pred. No. 20;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 4 PQOWFWL 11  
|:|:|:  
Db 86 PSRWYLL 93  
RESULT 41  
S76097  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S76097  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
S.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76097  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-208 <KAN>  
A:Cross-references: EMBL:D63999; GB:AB001339; NID:q1001396; PIDN:BAA10075.1; PID:g100  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Start codon: GTG

Query Match 54.9%; Score 39; DB 2; Length 208;  
Best Local Similarity 55.8%; Pred. No. 42;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 2 PKPQOWFWL 10  
|:|:|:  
Db 34 PSPQPQWMI 42  
RESULT 42  
B86503  
acyltransferase [imported] - Chlamydomophila pneumoniae (strain J138)  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001  
C:Accession: B86503  
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;  
Nucleic Acids Res. 28, 2311-2314, 2000  
A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138:  
A:Reference number: A86491; MUID:20330349  
A:Accession: B86503  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-341 <STO>  
A:Cross-references: GB:BA000008; NID:g8978471; PIDN:BAA98308.1; GSPDB:GN00142  
A:Experimental source: strain J138  
C:Genetics:  
A:Gene: htrB\_2

Query Match 54.9%; Score 39; DB 2; Length 341;  
Best Local Similarity 50.0%; Pred. No. 69;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 3 KPQOWFWL 10  
|:|:|:  
Db 180 QPEQWMI 187  
RESULT 43  
A36669  
galactoside 3(4)-L-fucosyltransferase (EC 2.4.1.65) - human  
N:Alternate names: alpha (1,3/1,4) fucosyltransferase; blood group Lewis alpha-4-fuco  
C:Species: Homo sapiens (man)  
C>Date: 12-Apr-1991 #sequence\_revision 12-Apr-1991 #text\_change 29-Sep-1999  
C:Accession: A36669; I39043; I39044; I39045; S12123  
R:Kukowska-Latallo, J.F.; Larsen, R.D.; Nair, R.P.; Lowe, J.B.  
Genes Dev. 4, 1288-1303, 1990  
A:Title: A cloned human cDNA determines expression of a mouse stage-specific embryoni  
A:Reference number: A36669; MUID:91032981  
A:Accession: A36669  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-361 <KUK>  
A:Cross-references: GB:X53578; NID:g28529; PIDN:CAA37641.1; PID:g28530  
R:Cameron, H.S.; Szczepaniak, D.; Weston, B.W.  
J. Biol. Chem. 270, 20112-20122, 1995  
A:Title: Expression of human chromosome 19p alpha(1,3)-fucosyltransferase genes in no  
A:Reference number: I39043; MUID:95378269



A:Accession: I39043  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-361 <RES>  
A:Cross-references: EMBL:U27326; NID:g967188; PIDN:AAC50185.1; PID:g967189  
A:Accession: I39044  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-361 <RE2>  
A:Cross-references: EMBL:U27327; NID:g967190; PIDN:AAC50186.1; PID:g967191  
A:Accession: I39045  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-361 <RE3>  
A:Cross-references: EMBL:U27328; NID:g967192; PIDN:AAC50187.1; PID:g967193  
C:Genetics:  
A:Gene: GDB:FUT3; LE  
A:Cross-references: GDB:135717; OMIM:111100  
A:Map position: 19p13.3-19p13.3  
A:Note: alternative splicing 5' to the coding region  
C:Superfamily: galactoside 3(4)-L-fucosyltransferase  
C:Keywords: glycosyltransferase; hexosyltransferase; transmembrane protein

Query Match 54.9%; Score 39; DB 2; Length 361;  
Best Local Similarity 55.6%; Pred. No. 73;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
||:|:|  
Db 126 RPQQRWIW 134

RESULT 44  
A42270  
alpha (1,3) fucosyltransferase FUT5 - human  
N:Alternate names: fucosyltransferase 5  
C:Species: Homo sapiens (man)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 29-Sep-1999  
C:Accession: A42270; I39046; I39047  
R:Weston, B.W.; Nair, R.P.; Larsen, R.D.; Lowe, J.B.  
J. Biol. Chem. 267, 4152-4160, 1992  
A:Title: Isolation of a novel human alpha (1,3)fucosyltransferase gene and molecular clones encoding enzymes with distinct acceptor substrate specificities.  
A:Reference number: A42270; MUID:92156161  
A:Accession: A42270  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-374 <WES>  
A:Cross-references: GB:M81485; NID:g128490; PIDN:AAA98117.1; PID:g1280209  
A:Note: sequence extracted from NCBI backbone (NCBIN:82825, NCBIIP:82826)  
R:Cameron, H.S.; Szczepaniak, D.; Weston, B.W.  
J. Biol. Chem. 270, 20112-20122; 1995  
A:Title: Expression of human chromosome 19p alpha(1,3)-fucosyltransferase genes in normal cells.  
A:Reference number: I39043; MUID:95378269  
A:Accession: I39046  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <RES>  
A:Cross-references: EMBL:U27329; NID:g967194; PIDN:AAC50188.1; PID:g967195  
A:Accession: I39047  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-374 <RE2>  
A:Cross-references: EMBL:U27330; NID:g967196; PIDN:AAC50189.1; PID:g967197  
C:Genetics:  
A:Gene: GDB:FUT5  
A:Cross-references: GDB:131644; OMIM:136835  
A:Map position: 19p13.3-19p13.3  
C:Superfamily: galactoside 3(4)-L-fucosyltransferase

Query Match 54.9%; Score 39; DB 2; Length 374;  
Best Local Similarity 55.6%; Pred. No. 73;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Best Local Similarity 55.6%; Pred. No. 76;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
||:|:|  
Db 139 RPQQRWIW 147

RESULT 45  
D71803  
ubiquinol--cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Helicobacter pylori (s)  
C:Species: Helicobacter pylori  
A:Variety: strain J99  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 20-Apr-2000  
C:Accession: D71803  
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.; Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.  
Nature 397, 176-180, 1999  
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric p  
A:Reference number: A71800; MUID:99120557  
A:Accession: D71803  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-412 <ARN>  
A:Cross-references: GB:AE001568; GB:AE001439; NID:g4156083; PIDN:AAD07046.1; PID:g415  
A:Experimental source: strain J99  
C:Genetics:  
A:Gene: petB  
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocou  
C:Keywords: electron transfer; heme; iron; metalloprotein; oxidoreductase  
F:21-370/Domain: cytochrome b homology <CVB>  
F:21-223/Domain: cytochrome b6 homology <CB6>  
F:250-370/Domain: plastocoulin--plastocyanin reductase 17K protein homology <17K>  
F:94,195/Binding site: heme iron (His) (axial ligands) (low potential) #status predic  
F:108,210/Binding site: heme iron (His) (axial ligands) (high potential) #status predic

Query Match 54.9%; Score 39; DB 2; Length 412;  
Best Local Similarity 54.5%; Pred. No. 84;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQQWF 11  
||:|:|  
Db 349 RPAFWFWLL 359

RESULT 46  
F83010  
probable oxidoreductase PA5084 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83010  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A:Reference number: AB2950; MUID:20437337  
A:Accession: F83010  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-416 <STO>  
A:Cross-references: GB:AE004921; GB:AE004091; NID:g9951372; PIDN:AAG08469.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA5084

Query Match 54.9%; Score 39; DB 2; Length 416;  
Best Local Similarity 53.8%; Pred. No. 84;  
Matches 7; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

QY 1 RPK--PQQWF 11

Db 80 RPLDPAQWRWLL 92  
||: | || ||:

## RESULT 47

B29413  
ubiquinol--cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Paracoccus denitrificans  
N:Alternate names: bcl complex; complex III  
C:Species: Paracoccus denitrificans  
C:Date: 31-Mar-1999 #sequence\_revision 20-Aug-1994 #text\_change 03-Mar-2000  
C:Accession: B29413  
R:Kuroski, B.; Ludwig, B.  
J. Biol. Chem. 262, 13805-13811, 1987  
A:Title: The genes of the Paracoccus denitrificans bc-1 complex. Nucleotide sequence and  
A:Reference number: A92613; MUID:88007612  
A:Accession: B29413  
A:Molecule type: DNA  
A:Residues: 1-440 <KUR>  
A:Cross-references: GB:M17522; NID:g150569; PIDN:AAA25572.1; PID:g150571  
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin  
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; oxidoreductase  
F:26-381/Domain: cytochrome b homology <CBH>  
F:26-225/Domain: cytochrome b6 homology <CB6>  
F:51-67/Domain: transmembrane #status predicted <TM1>  
F:96-114/Domain: transmembrane #status predicted <TM2>  
F:134-150/Domain: transmembrane #status predicted <TM3>  
F:195-217/Domain: transmembrane #status predicted <TM4>  
F:245-361/Domain: plastocyanin-plastocyanin reductase 17K protein homology <L7K>  
F:253-269/Domain: transmembrane #status predicted <TM5>  
F:330-346/Domain: transmembrane #status predicted <TM6>  
F:365-383/Domain: transmembrane #status predicted <TM7>  
F:395-411/Domain: transmembrane #status predicted <TM8>  
F:97,198/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted  
F:111,212/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 54.9%; Score 39; DB 1; Length 440;  
Best Local Similarity 54.5%; Pred. No. 89;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPKQOQFWLM 11  
||: ||||:  
Db 360 RPLFKWFWLL 370

## RESULT 48

F81436  
probable integral membrane protein with hemeolysin domain Cj0183 [imported] - Campylobacter  
C:Species: Campylobacter jejuni  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 21-Jul-2000  
C:Accession: F81436  
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Kettle, J.M.; Churcher, C.; Basham, D.; Chilling  
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barre  
Nature 403, 665-668, 2000  
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hyp  
A:Reference number: A81250; MUID:20150912  
A:Accession: F81436  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-452 <PAR>  
A:Cross-references: GB:AL111168; NID:g6967505; PIDN:CAB72666.1; PID:g696767  
A:Experimental source: serotype O2, strain NCTC 11168  
C:Genetics:  
A:Gene: Cj0183  
C:Superfamily: hypothetical protein HI0107

Query Match 54.9%; Score 39; DB 2; Length 452;  
Best Local Similarity 44.4%; Pred. No. 92;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPOQWFWLM 11  
||: ||||:

Db 160 RPLHFWWML 168

## RESULT 49

E81551  
lipid A biosynthesis lauroyl acyltransferase, probable CP0676 [imported] - Chlamydia  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
C:Accession: E81551  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke  
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe  
Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39  
A:Reference number: A81500; MUID:20150255  
A:Accession: E81551  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-462 <REA>  
A:Cross-references: GB:AE002225; GB:AE002161; NID:g7189583; PIDN:AAF38487.1; PID:g718  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: CP0676

Query Match 54.9%; Score 39; DB 2; Length 462;  
Best Local Similarity 50.0%; Pred. No. 94;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KPOQWFWL 10  
:|||:  
Db 301 QPEQWMI 308

## RESULT 50

B72119  
acyltransferase - Chlamydia pneumoniae (strain CWL029)  
C:Species: Chlamydia pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 05-May-2000  
C:Accession: B72119  
R:Kaiman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood,  
Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606  
A:Accession: B72119  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-467 <ARN>  
A:Cross-references: GB:AE001596; GB:AE001363; NID:g4376357; PIDN:AAD18251.1; PID:g437  
A:Experimental source: strain CWL029  
C:Genetics:  
A:Gene: htrB

Query Match 54.9%; Score 39; DB 2; Length 467;  
Best Local Similarity 50.0%; Pred. No. 95;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KPOQWFWL 10  
:|||:  
Db 301 QPEQWMI 308

## RESULT 51

E72682  
hypothetical protein APE0879 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000  
C:Accession: E72682  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Ta  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aero  
A:Reference number: A72450; MUID:99310339

A:Accession: E72682  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-480 <RAW>  
A:Cross-references: DDBJ:AF000060; NID:g5104188; PIDN:BAA79861.1; PID:g5104546  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE0879  
C:Superfamily: Aeropyrum pernix hypothetical protein APE0879

Query Match 54.9%; Score 39; DB 2; Length 480;  
Best Local Similarity 55.6%; Pred. No. 97;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQQWFWM 11  
||:||||:  
Db 283 KPQQWFMI 291

## RESULT 52

H71365

Probable linc protein (linc) - syphilis spirochete

C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)

C>Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999

C:Accession: H71365

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin  
rson, J.; Khaliak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo  
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

Science 281, 375-388, 1998  
A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.

A:Reference number: A71250; MUID:98332770

A:Accession: H71365

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-525 <COL>

A:Cross-references: GB:AF001195; GB:AF000520; NID:g3322366; PIDN:AAC26555.1; PID:g332237

A:Experimental source: strain Nichols

C:Genetics:

A:Gene: TP0107

Query Match 54.9%; Score 39; DB 2; Length 525;  
Best Local Similarity 55.6%; Pred. No. 1.1e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 9  
||:|:|:  
Db 173 RPNRANFW 181

## RESULT 53

TL3062

CLOCK protein - fruit fly (Drosophila melanogaster)

N:Alternate names: circadian rhythm protein

C:Species: Drosophila melanogaster

C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000

C:Accession: TL3062

R:Allada, R.; White, N.E.; So, W.V.; Hall, J.C.; Rosbash, M.

Cell 93, 791-804, 1998

A:Title: A mutant Drosophila homolog of mammalian CLOCK disrupts circadian rhythms and t

A:Reference number: 217596; MUID:98292177

A:Accession: TL3062

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1015 <ALL>

A:Cross-references: EMBL:AF065133; NID:g3213257; PID:g3213258; PIDN:AAC39101.1

C:Genetics:

A:Gene: Clk

A:Cross-references: FlyBase:FBgn0023076

A:Map position: 3

Query Match 54.9%; Score 39; DB 2; Length 1015;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQQWFML 10  
|:||||:  
Db 331 KGQQWIWL 338

## RESULT 54

TL3068

CLOCK protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000

C:Accession: TL3068

R:Darlington, T.K.; Wager-Smith, K.; Ceriani, M.F.; Staknis, D.; Gekakis, N.; Steeves  
Science 280, 1599-1603, 1998

A:Title: Closing the circadian loop: CLOCK-induced transcription of its own inhibitor  
A:Reference number: Z17599; MUID:98279147

A:Accession: TL3068

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1023 <DAR>

A:Cross-references: EMBL:AF067207; NID:g3192866; PID:g3192867; PIDN:AAD10630.1

C:Genetics:

A:Cross-references: FlyBase:FBgn0023076

C:Function:

A:Description: required for circadian behavioral rhythms

Query Match 54.9%; Score 39; DB 2; Length 1023;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQQWFML 10  
|:||||:  
Db 331 KGQQWIWL 338

## RESULT 55

TL3071

CLOCK protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C>Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000

C:Accession: TL3071

R:Bae, K.; Lee, C.; Sidote, D.; Chuang, K.Y.; Edery, I.

Mol. Cell. Biol. 18, 6142-6151, 1998

A:Title: Circadian regulation of a drosophila homolog of the mammalian clock gene: PE  
A:Reference number: Z17601; MUID:98414630

A:Accession: TL3071

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1027 <BAE>

A:Cross-references: EMBL:AF069997; NID:g3219725; PID:g3219726; PIDN:AACG2234.1

C:Genetics:

A:Cross-references: FlyBase:FBgn0023076

C:Function:

A:Description: required for circadian behavioral rhythms

Query Match 54.9%; Score 39; DB 2; Length 1027;  
Best Local Similarity 75.0%; Pred. No. 2.1e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 KPQQWFML 10  
|:||||:  
Db 335 KGQQWIWL 342

## RESULT 56

D83242

hypothetical protein PA3216 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa



## A30533

Lymphocyte-specific protein LSP1 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 07-May-1999  
C:Accession: A30533; PC1325  
R:Jongstra, J.; Tidmarsh, G.F.; Jongstra-Bilen, J.; Davis, M.M.  
J. Immunol. 141, 3999-4004, 1988  
A:Title: A new lymphocyte-specific gene which encodes a putative Ca(2+)-binding protein  
A:Reference number: A30533; MUID:89035543  
A:Accession: A30533  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-330 <JON>  
R:Matsumoto, N.; Toyoshima, S.; Osawa, T.  
J. Biochem. 113, 630-636, 1993  
A:Title: Characterization of the 50 kDa protein phosphorylated in concanavalin A-stimulated lymphocytes  
A:Reference number: PC1325; MUID:93340113  
A:Accession: PC1325  
A:Molecule type: protein  
A:Residues: 55-77, 'X', 79-81, 131-144, 'XX', 211-218, 'X', 220-229, 238-242, 'X', 244-252, 'S', 254-255  
C:Genetics:  
A:Gene: LSP1  
C:Keywords: calcium binding; phosphoprotein

Query Match 53.5%; Score 38; DB 2; Length 330;  
Best Local Similarity 55.6%; Pred. No. 96;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
:|:|:|  
DB 92 KPEPQQWF 100

## RESULT 61

Lymphocyte-specific protein - mouse (fragment)  
C:Species: Mus sp. (mouse)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 05-Nov-1999  
C:Accession: I57835  
R:Jongstra, J.; Ittel, M.E.; Iscove, N.N.; Brady, G.  
Mol. Immunol. 31, 1125-1131, 1994  
A:Title: The LSP1 gene is expressed in cultured normal and transformed mouse macrophages  
A:Reference number: I57835; MUID:95021322  
A:Accession: I57835  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-330 <RES>  
A:Cross-references: GB:S74179; NID:G709978; PIDN:AAB32257.1; PID:G709979  
C:Genetics:  
A:Gene: LSP1

Query Match 53.5%; Score 38; DB 2; Length 330;  
Best Local Similarity 55.6%; Pred. No. 96;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
:|:|:|  
DB 92 KPEPQQWF 100

## RESULT 62

JE0111  
Lectin-like oxidized LDL receptor - mouse  
N:Alternate names: LDX-1  
C:Species: Mus musculus (house mouse)  
C:Date: 22-May-1998 #sequence\_revision 29-May-1998 #text\_change 07-May-1999  
C:Accession: JE0111  
R:Hoshikawa, H.; Sawamura, T.; Kakutani, M.; Aoyama, T.; Nakamura, T.; Masaki, T.  
Biochem. Biophys. Res. Commun. 245, 841-846, 1998  
A:Title: High affinity binding of oxidized LDL to mouse lectin-like oxidized LDL receptor  
A:Reference number: JE0111; MUID:98249801

A:Accession: JE0111  
A:Molecule type: mRNA  
A:Residues: 1-363 <HOS>  
F:34-59/Domain: transmembrane #status predicted <TM>

Query Match 53.5%; Score 38; DB 2; Length 363;  
Best Local Similarity 62.5%; Pred. No. 11e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 RPKPQQWF 9  
:|:|:|  
DB 234 PCPQDWIW 241

## RESULT 63

B82697  
rod shape-determining protein Xfl1313 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 15-Sep-2000  
C:Accession: B82697  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequence  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82697  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-373 <SIM>  
A:Cross-references: GB:AE003964; GB:AE003849; NID:G9106300; PIDN:AAF84122.1; GSPDB:GN  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, D.M.; Carraro, D.M.; Carrier  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
M.; Tshuko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: Xfl1313  
C:Superfamily: rod shape-determining protein

Query Match 53.5%; Score 38; DB 2; Length 373;  
Best Local Similarity 62.5%; Pred. No. 11e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQQWFWM 11  
:|:|:|  
DB 201 PFSFWLL 208

## RESULT 64

JC4591  
alpha-1,3 fucosyltransferase (EC 2.4.1.-) - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 16-Apr-1996 #sequence\_revision 24-May-1996 #text\_change 24-Nov-1999  
C:Accession: JC4591  
R:Ozawa, M.; Muramatsu, T.  
J. Biochem. 119, 302-308, 1996  
A:Title: Molecular cloning and expression of a mouse alpha-1,3 fucosyltransferase gen  
A:Reference number: JC4591; MUID:97037075  
A:Accession: JC4591  
A:Molecule type: mRNA  
A:Residues: 1-400 <OZA>

A:Cross-references: DBJ:D63379  
A:Experimental source: Embryonal carcinoma F9 cells  
C:Superfamily: galactoside 3(4)-L-fucosyltransferase  
C:Keywords: glycoprotein; glycosyltransferase; hexosyltransferase; transmembrane protein  
F:1-23/Domain: intracellular #status predicted <INT>  
F:24-49/Domain: transmembrane #status predicted <TRM>  
F:84,185/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.5%; Score 38; DB 2; Length 400;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWFNL 10  
||| |  
DB 156 RPPGQRWVM 165

RESULT 65  
B36340  
N:Alternate names: CD15; ELAM-1 ligand fucosyltransferase (ELFT); FCT3A; FUC-TIV; myeloid  
C:Species: Homo sapiens (man)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: B36340; A36340; A40976; A41202  
R:Goelz, S.E.; Hession, C.; Goff, D.; Griffiths, B.; Tizard, R.; Newman, B.; Chi-Rosso, R.  
Cell 63, 1349-1356, 1990  
A:Title: ELFT: a gene that directs the expression of an ELAM-1 ligand.  
A:Reference number: A36340; MUID:91084863  
A:Accession: B36340  
A:Molecule type: mRNA  
A:Residues: 1-405 <GOE1>  
A:Cross-references: GB:M58596; NID:g182068; PIDN:AAA63172.1; PID:g182069  
A:Accession: A36340  
A:Molecule type: mRNA  
A:Residues: 'MRLWGARKSGAGWEKEAEQPCANSRLGPGR','SGRKRAVPGWASWPAHLALARPRLHGGAGC  
A:Cross-references: GB:M58597; NID:g182070; PIDN:AAA63173.1; PID:g182071  
A>Note: the codon used as an initiator for this translation is not in a good context for  
R:Low, J.B.; Kukowska-Latallo, J.F.; Nair, R.P.; Larsen, R.D.; Marks, R.M.; Macher, B.A.  
J. Biol. Chem. 266, 17467-17477, 1991  
A:Title: Molecular cloning of a human fucosyltransferase gene that determines expression  
A:Reference number: A40976; MUID:91373370  
A:Accession: A40976  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-86,'p',88-405 <LOW>  
A:Cross-references: GB:M65030; NID:g182791; PIDN:AAA92977.1; PID:g1236720  
R:Kumar, R.; Potvin, B.; Muller, W.A.; Stanley, P.  
J. Biol. Chem. 266, 21777-21783, 1991  
A:Title: Cloning of a human alpha(1,3)-fucosyltransferase gene that encodes ELFT but doe  
A:Reference number: A41202; MUID:92042084  
A:Accession: A41202  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-240,'D',242-400 <KUM>  
A:Cross-references: GB:S65161; NID:g239005; PIDN:AAB20349.1; PID:g239006  
C:Genetics:  
A:Gene: GDB:FUT4; CD15; FCT3A; FUC-TIV  
A:Cross-references: GDB:131563; OMIM:104230  
A:Map position: 11q21-11q21  
C:Superfamily: galactoside 3(4)-L-fucosyltransferase  
C:Keywords: glycoprotein; glycosyltransferase; hexosyltransferase  
F:1-48/Domain: signal sequence #status predicted <SIG>  
F:49-405/Product: alpha(1,3)-fucosyltransferase 4 #status predicted <MAT>  
F:91,190/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 53.5%; Score 38; DB 2; Length 405;  
Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWFNL 10  
||| |  
DB 156 RPPGQRWVM 165

Db 161 RPPGQRWVM 170

RESULT 66  
C64712  
ubiquinol--cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Helicobacter pylori (s  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 20-Apr-2000  
C:Accession: C64712  
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.  
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; Mcke  
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey,  
Nature 388, 539-547, 1997  
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,  
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.  
A:Reference number: A64520; MUID:97394467  
A:Accession: C64712  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-412 <TOM>  
A:Cross-references: GB:AE000652; GB:AE000511; NID:g2314720; PIDN:AAD08579.1; PID:g231  
C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoqui  
C:Keywords: electron transfer; heme; iron; metalloprotein; oxidoreductase  
F:21-370/Domain: cytochrome b6 homology <CBH>  
F:21-223/Domain: cytochrome b6 homology <CBH>  
F:250-370/Domain: plastocyanin reductase 17K protein homology <17K>  
F:94,195/Binding site: heme iron (His) (axial ligands) (low potential) #status predic  
F:108,210/Binding site: heme iron (His) (axial ligands) (high potential) #status pred

Query Match 53.5%; Score 38; DB 2; Length 412;  
Best Local Similarity 54.5%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQQWFNL 11  
||| |  
DB 349 RPAFMVFWLV 359

RESULT 67  
C96806  
unknown protein T5M16.25 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C96806  
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: C96806  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-421 <STO>  
A:Cross-references: GB:AE005173; NID:g6382510; PIDN:AAF07796.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T5M16.25  
A:Map position: 1

Query Match 53.5%; Score 38; DB 2; Length 421;  
Best Local Similarity 62.5%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWF 8  
: ||| |  
DB 145 KRPQVQWY 152

```
RESULT 68
A57396
alpha-1,3-fucosyltransferase Fuc-TIV (EC 2.4.1.1) - mouse
N:Alternate names: ELAM-1 ligand fucosyltransferase homolog
C:Species: Mus musculus (house mouse)
C:Date: 08-Feb-1996 #sequence_revision 08-Feb-1996 #text_change 11-Jan-2000
C:Accession: A57596
R:Gersten, K.M.; Natsuka, S.; Trinchera, M.; Petryniak, B.; Kelly, R.J.; Hiraiwa, N.; Je
J. Biol. Chem. 270, 25047-25056, 1995
A:Title: Molecular cloning, expression, chromosomal assignment, and tissue-specific exp
erise.
A:Reference number: A57596; MUID:96027607
A:Accession: A57596
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-433 <GER>
A:Cross-references: GB:U33457; NID:g1039426; PIDN:AAC52269.1; PID:g1039427
C:Superfamily: galactoside 3(4)-L-fucosyltransferase
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 53.5%; Score 38; DB 2; Length 433;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPQPQWFNL 10
|| |||
Db 189 RPPGQWVWM 198

RESULT 69
T04448
hypothetical protein F4D11.30 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 30-Apr-1999
C:Accession: T04448
R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoerge, W.; Hohelsel, J.; Mewes, H.W
submitted to the Protein Sequence Database, April 1998
A:Reference number: T15360
A:Accession: T04448
A:Molecule type: DNA
A:Residues: 1-455 <BEV>
A:Cross-references: EMBL:AL022537
A:Experimental source: cultivar Columbia; BAC clone F4D11
C:Genetics:
A:Map position: 4
A:Introns: 106/1; 132/3; 158/2; 215/1; 247/3; 283/3; 322/3; 372/3
A:Note: F4D11.30

Query Match 53.5%; Score 38; DB 2; Length 455;
Best Local Similarity 57.1%; Pred. No. 1.3e+02;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 4 PQQWFNL 10
|::|||
Db 242 PRKWFV 248

RESULT 70
S02489
hypothetical protein SPAC4G8.l2c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 16-May-1996 #sequence_revision 13-Mar-1997 #text_change 31-Jan-2000
C:Accession: T38857; S62489
R:Badcock, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.
submitted to the EMBL Data Library, October 1995
A:Reference number: T38857
A:Accession: T38857
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA

A:Residues: 1-533 <BA2>
A:Cross-references: EMBL:Z56276; NID:g1022345; PIDN:CAA91213.1; PID:g1022357; GSPDB:G
A:Experimental source: strain 972h-; cosmid c4G8
C:Genetics:
A:Gene: SPAC4G8.l2c
A:Map position: 1L
A:Introns: 42/1; 84/1; 107/1; 162/3; 215/2

Query Match 53.5%; Score 38; DB 2; Length 533;
Best Local Similarity 55.6%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 3 KPQQWFNL 11
|| |||
Db 296 KPATWLL 304

RESULT 71
I59550
aryl hydrocarbon receptor nuclear translocator Arnt [imported] - human
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
C:Accession: I59550
R:Hoffman, E.C.; Reyes, H.; Chu, F.F.; Sander, F.; Conley, L.H.; Brooks, B.A.; Hankin
Science 252, 954-958, 1991
A:Title: Cloning of a factor required for activity of the Ah (dioxin) receptor.
A:Reference number: I59550; MUID:91240280
A:Accession: I59550
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-789 <RES>
A:Cross-references: GB:M69238; NID:g179003; PIDN:AAA51777.1; PID:g179004
C:Genetics:
A:Gene: GDB:ARNT
A:Cross-references: GDB:119701; OMIM:126110
A:Map position: 1q21-1q21

Query Match 53.5%; Score 38; DB 2; Length 789;
Best Local Similarity 50.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 RPQPQWFNL 10
|| |||
Db 430 RSKNQEWLWM 439

RESULT 72
WMNVN
104K glycoprotein - Trichoplusia ni granulosis virus
C:Species: Trichoplusia ni granulosis virus, TnGV.
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jun-2000
C:Accession: J01328
R:Hashimoto, Y.; Corsaro, B.G.; Granados, R.R.
J. Gen. Virol. 72, 2645-2651, 1991
A:Title: Location and nucleotide sequence of the gene encoding the viral enhancing fa
A:Reference number: JQ1328; MUID:92044434
A:Accession: JQ1328
A:Molecule type: DNA
A:Residues: 1-901 <HAS>
A:Cross-references: GB:D12617; NID:g221443; PIDN:BAA02141.1; PID:g221444
A:Note: the authors translated the codon CTA for residue 12 as Val
C:Comment: This protein is involved in disruption of the peritrophic membrane and fus
C:Superfamily: Trichoplusia ni granulosis virus 104K glycoprotein
C:Keywords: glycoprotein
F:65,265,306,339,349,540,594,595,621,642,683,698/Binding site: carbohydrate (Asn) (CO

Query Match 53.5%; Score 38; DB 1; Length 901;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

QY 2 PKPQOWFWL 10  
| | | | |  
Db 353 PYPQIWSWL 361

## RESULT 73

S76899  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76899  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis s.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76899  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-909 <KAN>  
A:Cross-references: EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BAAL8811.1; PID:dl01954  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Start codon: GTG

Query Match 53.5%; Score 38; DB 2; Length 909;  
Best Local Similarity 71.4%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 PKPQOWF 8  
| | | | |  
Db 445 PRPQSWF 451

## RESULT 74

T09484  
cartilage intermediate layer protein precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T09484  
R:Lorenz, P.; Neame, P.; Sommarin, Y.; Heinegard, D.  
J. Biol. Chem. 273, 23469-23475, 1998  
A:Title: Cloning and deduced amino acid sequence of a novel cartilage protein (CILP) id  
A:Reference number: Z16689; MUID:98389785  
A:Accession: T09484  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1184 <LOR>  
A:Cross-references: EMBL:AF035408; NID:g3513502; PIDN:AAC33838.1; PID:g3513503  
A:Experimental source: tissue type articular cartilage  
C:Genetics:  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-1184/Product: cartilage intermediate layer protein #status predicted <MAT>

Query Match 53.5%; Score 38; DB 2; Length 1184;  
Best Local Similarity 44.4%; Pred. No. 3.4e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQOWFW 9  
: | | : | |  
Db 335 KPRPKDKYFW 343

## RESULT 75

T31113  
mucin-like glycoprotein 900 - Cryptosporidium parvum  
C:Species: Cryptosporidium parvum  
C:Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999

C:Accession: T31113  
R:Barnes, D.A.; Bonnin, A.; Huang, J.X.; Gousset, L.; Wu, J.; Gut, J.; Doyle, P.; Dub Mol. Biochem. Parasitol. 96, 93-110, 1998  
A:Title: A novel multi-domain mucin-like glycoprotein of Cryptosporidium parvum media  
A:Reference number: Z20989; MUID:99066935

A:Accession: T31113  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1832 <BAR>  
A:Cross-references: EMBL:AF068065; NID:g4063041; PID:g4063042; PIDN:AAC98153.1

Query Match 53.5%; Score 38; DB 2; Length 1832;  
Best Local Similarity 62.5%; Pred. No. 5.2e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQOWFWL 10  
| | : | | |  
Db 527 KPDEWCWL 534

## RESULT 76

D69351  
hypothetical protein AF0812 - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 04-Mar-2000  
C:Accession: D69351  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997  
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
A:Reference number: A69250; MUID:98049343  
A:Accession: D69351  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-298 <KLE>  
A:Cross-references: GB:AE001048; GB:AE000782; NID:g2689371; PIDN:AAB90432.1; PID:g264  
C:Superfamily: Archaeoglobus fulgidus hypothetical protein AF0812

Query Match 52.8%; Score 37.5; DB 2; Length 293;  
Best Local Similarity 66.7%; Pred. No. 1e+02; 1; Indels 1; Gaps 1;  
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 2 PKPQOWF-W 9  
| | | | |  
Db 62 PKPEYWRW 70

## RESULT 77

JQ0846  
DNA-binding protein - equine herpesvirus 1 (fragment)  
C:Species: equine herpesvirus 1  
C:Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 23-Feb-1997  
C:Accession: JQ0846  
R:Bell, C.W.; Whalley, J.M.  
submitted to JIPID, January 1991  
A:Reference number: JQ0846  
A:Accession: JQ0846  
A:Molecule type: DNA  
A:Residues: 1-375 <BEL>  
C:Superfamily: herpesvirus DNA-binding protein  
C:Keywords: DNA binding; nucleus

Query Match 52.8%; Score 37.5; DB 2; Length 375;  
Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PKPQOWFWL 11



Db 10 PNP-QWFWTL 18  
| | | | |

## RESULT 78

DNBEV1  
DNA-binding protein UL29 - human herpesvirus 1 (strain 17)  
C:Species: human herpesvirus 1  
A:Note: host Homo sapiens (man)  
C:Date: 04-Dec-1986 #sequence\_revision 04-Dec-1986 #text\_change 16-Jun-2000  
C:Accession: A03790; B30085  
R:Quinn, J.P.; McGeoch, D.J.  
Nucleic Acids Res. 13, 8143-8163, 1985  
A:Title: DNA sequence of the region in the genome of herpes simplex virus type 1 containing  
A:Reference number: A93601; MUID:86067223  
A:Accession: A03790  
A:Molecule type: DNA  
A:Residues: 1-1196 <QUI>  
A:Cross-references: GB:M12356; PIDN:CAA26940.1; PID:g59863  
A:Experimental source: strain 17  
R:McGeoch, D.J.; Balrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perz  
J. Gen. Virol. 69, 1531-1574, 1988  
A:Title: The complete DNA sequence of the long unique region in the genome of herpes sim  
A:Reference number: A30083; MUID:88274327  
A:Accession: B30085  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-1196 <MCG>  
A:Cross-references: GB:D10879; PIDN:BAA01675.1; PID:g221750; GB:D00317  
C:Genetics:  
A:Gene: UL29  
A:Map position: 0.38-0.409  
C:Superfamily: herpesvirus DNA-binding protein  
C:Keywords: DNA binding

Query Match 52.8%; Score 37.5; DB 1; Length 1196;  
Best Local Similarity 66.7%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPKPOQWFW 9  
: | | | | |

Db 837 QPNP-QWFW 844

## RESULT 79

DNBEK5  
DNA-binding protein - human herpesvirus 1 (strain KOS1.1)  
C:Species: human herpesvirus 1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 16-Jul-1999  
C:Accession: A28601  
R:Gao, M.; Bouchev, J.; Curtin, K.; Knipe, D.M.  
Virology 163, 319-329, 1988  
A:Title: Genetic identification of a portion of the herpes simplex virus ICP8 protein re  
A:Reference number: A28601; MUID:88179536  
A:Accession: A28601  
A:Molecule type: DNA  
A:Residues: 1-1196 <GAO>  
A:Cross-references: GB:M20165; PIDN:AAA45793.1; PID:g330121  
C:Genetics:  
A:Map position: 0.38-0.409  
C:Superfamily: herpesvirus DNA-binding protein  
C:Keywords: DNA binding

Query Match 52.8%; Score 37.5; DB 1; Length 1196;  
Best Local Similarity 66.7%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPKPOQWFW 9  
: | | | | |

Db 837 QPNP-QWFW 844

## RESULT 80

DNBEHF  
DNA-binding protein - human herpesvirus 1 (strain F)  
C:Species: human herpesvirus 1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 07-Jun-1996  
C:Accession: D29242  
R:Hammerschmidt, W.; Conraths, F.; Mankertz, J.; Pauli, G.; Ludwig, H.; Buhk, H.J.  
Virology 165, 388-405, 1988  
A:Title: Conservation of a gene cluster including glycoprotein B in bovine herpesviro  
A:Reference number: A94381; MUID:88306231  
A:Accession: D29242  
A:Molecule type: DNA  
A:Residues: 1-1196 <HAM>  
A:Cross-references: GB:M21629  
C:Superfamily: herpesvirus DNA-binding protein  
C:Keywords: DNA binding

Query Match 52.8%; Score 37.5; DB 1; Length 1196;  
Best Local Similarity 66.7%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPKPOQWFW 9  
: | | | | |

Db 837 QPNP-QWFW 844

## RESULT 81

A48350  
DNA-binding protein - human herpesvirus 2  
C:Species: human herpesvirus 2  
A:Note: host Homo sapiens (man)  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 31-May-1996  
C:Accession: A48350  
R:Toh, Y.; Liu, Y.; Tanaka, S.; Mori, R.  
Arch. Virol. 129, 183-196, 1993  
A:Title: Nucleotide sequence of the major DNA-binding protein gene of herpes simplex  
A:Reference number: A48350; MUID:93228441  
A:Accession: A48350  
A:Molecule type: DNA  
A:Residues: 1-1197 <TOH>  
A:Note: sequence extracted from NCBI backbone (NCBIN:129069, NCBI:129070)  
C:Genetics:  
A:Map position: 0.375-0.405  
C:Superfamily: herpesvirus DNA-binding protein  
C:Keywords: DNA binding; zinc finger  
F:493-512/Region: zinc finger

Query Match 52.8%; Score 37.5; DB 1; Length 1197;  
Best Local Similarity 66.7%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 RPKPOQWFW 9  
: | | | | |

Db 837 QPNP-QWFW 844

## RESULT 82

DNBE29  
DNA-binding protein - human herpesvirus 3  
C:Species: human herpesvirus 3, varicella-zoster virus  
C:Date: 30-Sep-1988 #sequence\_revision 30-Sep-1988 #text\_change 16-Jul-1999  
C:Accession: C27214  
R:Davidson, A.J.; Scott, J.E.  
J. Gen. Virol. 67, 1759-1816, 1986  
A:Title: The complete DNA sequence of varicella-zoster virus.  
A:Reference number: A27345; MUID:86306657  
A:Accession: C27214  
A:Molecule type: DNA

A;Residues: 1-1204 <DAV>  
A;Cross-references: EMBL:X04370; NID:g59989; PIDN:CAA27912.1; PID:g60018  
C;Genetics:  
A;Gene: 29  
C;Superfamily: herpesvirus DNA-binding protein  
C;Keywords: DNA binding

Query Match 52.8%; Score 37.5; DB 1; Length 1204;  
Best Local Similarity 60.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PKPOQWFWM 11  
| | | | |  
Db 836 PNP-QWFWTL 844

RESULT 83

T42574  
DNA-binding protein - equine herpesvirus 4 (strain NS80567)  
C;Species: equine herpesvirus 4  
A;Variety: strain NS80567  
C;Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 21-Jul-2000  
C;Accession: T42574  
R;Telford, E.A.; Watson, M.S.; Perry, J.; Cullinane, A.A.; Davison, A.J.  
J. Gen. Virol. 79, 1197-1203, 1998  
A;Title: The DNA sequence of equine herpesvirus-4.  
A;Reference number: 222173; MUID:98264497  
A;Accession: T42574  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-1208 <TEL>  
A;Cross-references: EMBL:AF030027; NID:g2605950; PIDN:AAC59547.1; PID:g2605975  
A;Experimental source: strain NS80567  
C;Genetics:  
A;Gene: 31  
C;Superfamily: herpesvirus DNA-binding protein  
C;Keywords: DNA binding

Query Match 52.8%; Score 37.5; DB 2; Length 1208;  
Best Local Similarity 60.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PKPOQWFWM 11  
| | | | |  
Db 844 PNP-QWFWTL 852

RESULT 84

DNBEC4  
DNA-binding protein - equine herpesvirus 1 (strain AB4p)  
C;Species: equine herpesvirus 1  
A;Note: host Equus caballus (domestic horse)  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 16-Jul-1999  
C;Accession: E36798  
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.  
submitted to GenBank, March 1992  
A;Description: The DNA sequence of equine herpesvirus-1.  
A;Reference number: A36805  
A;Accession: E36798  
A;Molecule type: DNA  
A;Residues: 1-1209 <TEL>  
A;Cross-references: GB:M86664; NID:g330791; PIDN:AAB02466.1; PID:g330823  
R;Telford, E.A.R.; Watson, M.S.; McBride, K.; Davison, A.J.  
Virology 189, 304-316, 1992  
A;Title: The DNA sequence of equine herpesvirus-1.  
A;Reference number: A41831; MUID:92295566  
A;Contents: annotation; possible protein-coding frames  
A;Note: neither amino acid nor nucleotide sequence is given  
C;Genetics:  
A;Gene: 31  
C;Superfamily: herpesvirus DNA-binding protein

C;Keywords: DNA binding

Query Match 52.8%; Score 37.5; DB 1; Length 1209;  
Best Local Similarity 60.0%; Pred. No. 4.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 2 PKPOQWFWM 11  
| | | | |  
Db 844 PNP-QWFWTL 852

RESULT 85

S23308  
substance P - rainbow trout  
C;Species: Oncorhynchus mykiss (rainbow trout)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 18-Aug-2000  
C;Accession: S23308  
R;Jensen, J.; Conlon, J.M.  
Eur. J. Biochem. 206, 659-664, 1992  
A;Title: Substance-P-related and neurokinin-A-related peptides from the brain of the  
A;Reference number: S23186; MUID:92298992  
A;Accession: S23308  
A;Molecule type: protein  
A;Residues: 1-11 <JEN>  
A;Experimental source: brain  
C;Function:  
A;Description: may play a physiological role in the regulation of cardiovascular and  
A;Note: substance P is derived by post-translational processing of preprotachykinin A  
C;Superfamily: unassigned animal peptides  
C;Keywords: neuropeptide; amidated carboxyl end; tachykinin  
F;11/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 52.1%; Score 37; DB 2; Length 11;  
Best Local Similarity 54.5%; Pred. No. 4.7;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQWFWM 11  
: | : | : | : |  
Db 1 KPRPHQFFGLM 11

RESULT 86

S50061  
DNA binding protein, replication-origin specific - Escherichia coli  
C;Species: Escherichia coli  
C;Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 08-Oct-1999  
C;Accession: S50061; S37475  
R;Herrick, J.; Kern, R.; Guha, S.; Landoulsi, A.; Fayet, O.; Malki, A.; Kohiyama, M.  
EMBO J. 13, 4695-4703, 1994  
A;Title: Parental strand recognition of the DNA replication origin by the outer membr  
A;Reference number: S50061; MUID:95009972  
A;Accession: S50061  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-175 <HER>  
A;Cross-references: EMBL:226592  
R;Herrick, J.; Kern, R.; Fayet, O.; Guha, S.; Kohiyama, M.  
submitted to the EMBL data Library, September 1993  
A;Description: Cloning and characterization of a novel DNA binding protein specific o  
A;Reference number: S37475  
A;Accession: S37475  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 85-97, 'A', 99-149, 'ACEAIFNCLDRNPMNGLFSKPGNVSPPVVR' <HE2>  
A;Cross-references: EMBL:226592; NID:g404853; PIDN:CAA81346.1; PID:g404854

Query Match 52.1%; Score 37; DB 2; Length 175;  
Best Local Similarity 71.4%; Pred. No. 73;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPOOW 7  
||||:|  
Db 160 RPKPNEW 166

## RESULT 87

hypothetical protein SPAC7D4.09c - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T39087  
R:Gentles, S.; Churcher, C.M.; Wood, V.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, September 1997  
A:Reference number: 221826  
A:Accession: T39087  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-274 <GEN>  
A:Cross-references: EMBL:Z99532; PIDN:CAB16726.2; GSPDB:GNO0066; SPDB:SPAC7D4.09c  
A:Experimental source: strain 972h-; cosmid c7D4  
C:Genetics:  
A:Gene: SPDB:SPAC7D4.09c  
A:Map position: 1

Query Match 52.1%; Score 37; DB 2; Length 274;  
Best Local Similarity 66.7%; Pred. No. 1.1e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 4 PQOWFW 9  
|::|||  
Db 61 PKRWF 66

## RESULT 88

alpha(1,3)-fucosyltransferase (EC 2.4.1.-) 7 precursor - human  
N:Alternate names: leukocyte fucosyltransferase FucTVII  
C:Species: Homo sapiens (man)  
C:Date: 02-Aug-1994 #sequence\_revision 02-Aug-1994 #text\_change 20-Apr-2000  
C:Accession: A54057; A53713  
R:Sasaki, K.; Kurata, K.; Funayama, K.; Nagata, M.; Watanabe, E.; Ohta, S.; Hanai, N.; N  
J. Biol. Chem. 269, 14730-14737, 1994  
A:Title: Expression cloning of a novel alpha1,3-fucosyltransferase that is involved in b  
A:Reference number: A54057; MUID:94237894  
A:Accession: A54057  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-342 <SAS>  
A:Cross-references: GB:X78031; NID:g516292; PIDN:CAA54962.1; PID:g516293  
R:Natsuka, S.; Gersten, K.M.; Zenita, K.; Kannagi, R.; Lowe, J.B.  
J. Biol. Chem. 269, 16789-16794, 1994  
A:Title: Molecular cloning of a cDNA encoding a novel human leukocyte alpha-1,3-fucosylt  
A:Reference number: A53713; MUID:94266898  
A:Accession: A53713  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-160,'A',163-303,'SV',306-342 <NAT>  
A:Cross-references: GB:U08112; NID:g520463; PIDN:AAA56869.1; PID:g520464  
C:Genetics:  
A:Gene: GDB:FUT7  
A:Cross-references: GDB:373982  
A:Map position: 9pter-9qter  
C:Superfamily: galactoside 3(4)-L-fucosyltransferase  
C:Keywords: glycoprotein; glycosyltransferase; hexosyltransferase  
F:1-34/Domain: signal sequence #status predicted <SIG>  
F:35-342/Product: alpha(1,3)-fucosyltransferase 7 #status predicted <MAT>  
F:81,291/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 52.1%; Score 37; DB 2; Length 342;  
Best Local Similarity 55.6%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPOOWFW 9  
||:| | |  
Db 110 RPRGPWF 118

## RESULT 89

alpha(1,3/4)-fucosyltransferase - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 01-Aug-1995 #sequence\_revision 01-Sep-1995 #text\_change 13-Sep-1998  
C:Accession: S55498  
R:Oulmouden, A.; Wierinckx, A.; Petit, J.M.; Julien, R.  
submitted to the EMBL Data Library, June 1995  
A:Description: Molecular cloning and expression of bovine alpha (1,3/4)-fucosyltransf  
A:Reference number: S55498  
A:Accession: S55498  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-365 <OUL>  
A:Cross-references: EMBL:X87810; NID:g860807; PID:g860808  
C:Superfamily: galactoside 3(4)-L-fucosyltransferase

Query Match 52.1%; Score 37; DB 2; Length 365;  
Best Local Similarity 55.6%; Pred. No. 1.5e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPOOWFW 9  
|| | | |  
Db 130 RPDQRWF 138

## RESULT 90

H64861  
hypothetical protein b1163 - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999  
C:Accession: H64861  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: H64861  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-403 <BLAT>  
A:Cross-references: GB:AE000215; GB:U00096; NID:gl787405; PIDN:AAAC74247.1; PID:gl7874  
A:Experimental source: strain K-12, substrain MG1655

Query Match 52.1%; Score 37; DB 2; Length 403;  
Best Local Similarity 57.1%; Pred. No. 1.7e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQOWFWL 10  
|::| | |  
Db 367 PEEWML 373

## RESULT 91

S25953  
ubiquinol--cytochrome-c reductase (EC 1.10.2.2) cytochrome b - liverwort (Marchantia  
C:Species: Marchantia polymorpha  
C:Date: 07-May-1993 #sequence\_revision 02-Aug-1994 #text\_change 03-Mar-2000  
C:Accession: S25953  
R:Oda, K.; Yamato, K.; Ohta, E.; Nakamura, Y.; Takemura, M.; Nozato, N.; Akashi, K.;  
J. Mol. Biol. 223, 1-7, 1992  
A:Title: Gene organization deduced from the complete sequence of liverwort Marchantia  
A:Reference number: S25941; MUID:92114051  
A:Accession: S25953  
A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA  
A:Residues: 1-404 <ODA>

C:Cross-references: EMBL:M68929; NID:g786182; PIDN:AA09441.1; PID:g786227

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992

C:Genetics:

A:Gene: cob

A:Genome: mitochondrion  
A:Introns: 124/3; 261/3; 275/2

C:Function:

A:Description: the net reaction catalyzed by the ubiquinol--cytochrome-c reductase complex with two hydrogen ions taken up from the mitochondrial matrix and four hydrogen ions released into the intermembrane space

C:Pathway: oxidative phosphorylation; respiratory chain

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastocyanin

C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;

F:13-343/Domain: cytochrome b homology <CBH>

F:13-213/Domain: cytochrome b6 homology <CB6>

F:38-54/Domain: transmembrane #status predicted <TM1>

F:83-101/Domain: transmembrane #status predicted <TM2>

F:121-137/Domain: transmembrane #status predicted <TM3>

F:182-204/Domain: transmembrane #status predicted <TM4>

F:225-343/Domain: plastocyanin-plastocyanin reductase 17K protein homology <17K>

F:233-249/Domain: transmembrane #status predicted <TM5>

F:232-308/Domain: transmembrane #status predicted <TM6>

F:327-347/Domain: transmembrane #status predicted <TM7>

F:357-372/Domain: transmembrane #status predicted <TM8>

F:85,186/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:99,200/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 52.1%; Score 37; DB 1; Length 404;

Best Local Similarity 54.5%; Pred. No. 1.7e+02;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKPQQHFWLM 11

II I I I I I I

Db 322 RPIHQKFWLL 332

RESULT 92

B82147 conserved hypothetical protein VC1884 [imported] - Vibrio cholerae (strain N16961 serotype O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: B82147

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, R.

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: B82147

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-406 <HBI>

A:Cross-references: GB:AE004263; GB:AE003852; NID:g9656399; PIDN:AAF95032.1; GSPDB:GN001

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VC1884

A:Map position: 1

C:Superfamily: hypothetical protein HI1555

Query Match 52.1%; Score 37; DB 2; Length 406;

Best Local Similarity 55.6%; Pred. No. 1.7e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQHFW 9

I I I I I I

Db 243 QPLPQDMQW 251

RESULT 93

C82433

methyl-accepting chemotaxis protein VCA0658 [imported] - Vibrio cholerae (strain N169 C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: C82433

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers

l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: C82433

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-536 <HEI>

A:Cross-references: GB:AE004395; GB:AE003853; NID:g9658068; PIDN:AAF96559.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0658

A:Map position: 2

Query Match 52.1%; Score 37; DB 2; Length 536;

Best Local Similarity 71.4%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQWFWL 10

I I I I I I

Db 188 PQQWFWL 194

```
Query Match          52.1%; Score 37; DB 1; Length 655;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QQWFWL 10
    |||:|
Db 625 QQWWWL 630

RESULT 95
TOBPU
transposase - phage Mu
C:Species: phage Mu
C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 23-Jul-1999
C:Accession: A24746; S09550; S06975; S57318
R:Harshey, R.M.; Getroff, E.D.; Baldwin, D.L.; Miller, J.L.; Chaconas, G.
Proc. Natl. Acad. Sci. U.S.A. 82, 7676-7680, 1985
A:Title: Primary structure of phage mu transposase: homology to mu repressor.
A:Reference number: A24746; MUID:86067968
A:Accession: A24746
A:Molecule type: DNA
A:Residues: 1-662 <HAR>
A:Cross-references: GB:M11195
R:Wu, Z.; Chaconas, G.
EMBO J. 14, 3835-3843, 1995
A:Title: A novel DNA binding and nuclease activity in domain III of Mu transposase: evic
A:Reference number: S57318; MUID:95369255
A:Contents: annotation
R:Priess, H.; Kamp, D.; Kahmann, R.; Braeuer, B.; Delius, H.
Mol. Gen. Genet. 186, 315-321, 1982
A:Title: Nucleotide sequence of the immunity region of bacteriophage Mu.
A:Reference number: S07291; MUID:83012203
A:Accession: S09550
A:Molecule type: DNA
A:Residues: 1-88 <PRI>
A:Cross-references: EMBL:V01464; NID:g15807; PIDN:CAA24713.1; PID:g15810
C:Genetics:
A:Gene: A
C:Function:
A:Description: it is essential for integration, replication-transposition, and excision
C:Superfamily: phage Mu transposase
C:Keywords: DNA binding; DNA replication

Query Match          52.1%; Score 37; DB 1; Length 662;
Best Local Similarity 57.1%; Pred. No. 2.7e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KPQQWF 9
    |||:|
Db 285 RPKTWF 291

RESULT 96
I78558
hypothetical Brachyury (T)-related transcription factor - mouse
C:Species: Mus sp. (mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
C:Accession: I78558
R:Bulfone, A.; Smiga, S.M.; Shimamura, K.; Peterson, A.; Puellas, L.; Rubenstein, J.L.
Neuron 15, 63-78, 1995
A:Title: T-brain-1: a homolog of Brachyury whose expression defines molecularly distinct
A:Reference number: I58171; MUID:95344783
A:Accession: I78558
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-681 <RES>
A:Cross-references: GB:S78858; NID:g1222544; PIDN:AAA92011.1; PID:g1222545
C:Genetics:
A:Gene: Tbr-1/Tes-56
C:Superfamily: T-box homology
F:213-401/Domain: T-box homology <TBX>

Query Match          52.1%; Score 37; DB 2; Length 747;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PKPQWF 7
    |||:|
Db 528 PKPQWF 533

RESULT 97
I39444
AMP deaminase (EC 3.5.4.6) - human
N:Alternate names: myoadenylate deaminase
C:Species: Homo sapiens (man)
C:Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 22-Jun-1999
C:Accession: I39444
R:Sabina, R.L.; Fishbein, W.N.; Pezeshkpour, G.; Clarke, P.R.; Holmes, E.W.
Neurology 42, 170-179, 1992
A:Title: Molecular analysis of the myoadenylate deaminase deficiencies.
A:Reference number: I39444; MUID:92131279
A:Accession: I39444
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-747 <RES>
A:Cross-references: GB:M60092; NID:g178543; PIDN:AAA57281.1; PID:g178544
C:Genetics:
A:Gene: GDB:AMPD1
A:Cross-references: GDB:I19677; OMIM:102770
A:Map position: lp13-lp13
C:Superfamily: AMP deaminase
C:Keywords: hydrolase

Query Match          52.1%; Score 37; DB 2; Length 747;
Best Local Similarity 83.3%; Pred. No. 3.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 PKPQWF 7
    |||:|
Db 528 PKPQWF 533

RESULT 98
S54252
deep orange protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 24-Sep-1998
C:Accession: S54252
R:Shestopal, S.A.
submitted to the EMBL Data Library, April 1995
A:Reference number: S54252
A:Accession: S54252
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1002 <SHE>
A:Cross-references: EMBL:X86683; NID:g798831; PID:g798832
C:Genetics:
A:Gene: FlyBase:dor
A:Cross-references: FlyBase:FBgn0000482

Query Match          52.1%; Score 37; DB 2; Length 1002;
Best Local Similarity 71.4%; Pred. No. 4.1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 PQQFWL 10
    |||:|
Db 301 PKQAWL 307

RESULT 99
G02750
```

DNA-directed DNA polymerase (EC 2.7.7.7) gamma - human

C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 20-Sep-1999  
C:Accession: G02750

R:Ropp, P.A.; Copeland, W.C.

submitted to the EMBL Data Library, June 1996

A:Reference number: H01679

A:Accession: G02750

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1239 <ROP>

A:Cross-references: EMBL:U60325; NID:gl1399955; PID:gl1399956

C:Superfamily: DNA-directed DNA polymerase, mitochondrial

C:Keywords: nucleotidyltransferase

Query Match 52.1%; Score 37; DB 2; Length 1239;

Best Local Similarity 62.5%; Pred. No. 5e+02;

Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQWF 9

Db 164 PKPPAW 171

RESULT 100

T26656

hypothetical protein Y38E10A.f - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T26656

R:Wallis, J.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z20252

A:Accession: T26656

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1384 <WIL>

A:Cross-references: EMBL:AL110484; NID:el542205; PIDN:CAB54397.1; CESP:Y38E10A.f

A:Experimental source: clone Y38E10A

C:Genetics:

A:Gene: CESP:Y38E10A.f

A:Introns: 84/3; 115/3; 154/1; 187/3; 245/3; 325/3; 365/3; 422/3; 480/3; 525/3; 565/3; 6

Query Match 52.1%; Score 37; DB 2; Length 1384;

Best Local Similarity 54.5%; Pred. No. 5.6e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 1 RPKP--QQWF 9

Db 361 RPRPFVLKWF 371

RESULT 101

T32452

hypothetical protein F48A11.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999

C:Accession: T32452

R:Bradshaw, H.

submitted to the EMBL Data Library, September 1997

A:Description: The sequence of C. elegans cosmid F48A11.

A:Reference number: Z21171

A:Accession: T32452

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1635 <BRA>

A:Cross-references: EMBL:AF026210; PIDN:AAB71283.1; GSPDB:GNO0020; CESP:F48A11.1

A:Experimental source: strain Bristol N2, clone F48A11

C:Genetics:

A:Gene: CESP:F48A11.1

A:Map position: 2

A;Introns: 84/2; 108/1; 198/2; 293/3; 336/2; 477/3; 560/3; 612/3; 719/2; 952/2; 1371/

Query Match 52.1%; Score 37; DB 2; Length 1635;

Best Local Similarity 71.4%; Pred. No. 6.6e+02;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PQQWF 10

Db 580 PQTW 586

RESULT 102

T17420

probable polyketide synthase type I - Pseudomonas fluorescens

C:Species: Pseudomonas fluorescens

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Nov-2000

C:Accession: T17420

R:Nowak-Thompson, B.; Chaney, N.; Wing, J.S.; Gould, S.J.; Lopez, J.E.

J. Bacteriol. 181, 2166-2174, 1999

A:Title: Characterization of the pyoluteorin biosynthetic gene cluster of Pseudomonas

A:Reference number: Z18776; MUID:99194726

A:Accession: T17420

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-2458 <NOW>

A:Cross-references: EMBL:AF081920; NID:g4582974; PID:g2781416; PIDN:AAC38075.1

C:Genetics:

A:Gene: plbB

C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier pro

C:Keywords: carrier protein

F;31-429/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>

F;535-815/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT>

F;939-1009/Domain: acyl carrier protein homology <ACP1>

F;1053-1446/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>

F;2337-2408/Domain: acyl carrier protein homology <ACP2>

Query Match 52.1%; Score 37; DB 2; Length 2458;

Best Local Similarity 55.6%; Pred. No. 9.9e+02;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RRPQQWF 9

Db 1613 RPAGQLWY 1621

RESULT 103

JC4534

cytochrome P450 4F6 protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 15-Feb-1996 #sequence\_revision 19-Apr-1996 #text\_change 28-Jul-2000

C:Accession: JC4534

R:Kawashima, H.; Strobel, H.W.

Biochem. Biophys. Res. Commun. 217, 1137-1144, 1995

A:Title: cDNA cloning of three new forms of rat brain cytochrome P450 belonging to th

A:Reference number: JC4532; MUID:96125358

A:Accession: JC4534

A:Molecule type: mRNA

A:Residues: 1-537 <KAW>

A:Cross-references: NID:gl146439; PIDN:AAC52360.1; PID:gl146440

C:Experimental source: brain

C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology

C:Keywords: brain; chromoprotein; heme; iron; metalloprotein

F;325-490/Domain: cytochrome P450 homology <P45>

F;468/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 51.4%; Score 36.5; DB 2; Length 537;

Best Local Similarity 75.0%; Pred. No. 2.6e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2 PKPQQWF 9

Db 55 PKP-SWFW 61  
||| |||

## RESULT 104

S77658

hypothetical protein o659 - Mycobacterium leprae

C:Species: Mycobacterium leprae

C:Date: 11-Oct-1997 #sequence\_revision 24-Oct-1997 #text\_change 22-Oct-1999

C:Accession: S77658; S49520

R:Psihi, H.; Cole, S.T.

Mol. Microbiol. 16, 909-919, 1995

A:Title: The Mycobacterium leprae genome: systematic sequence analysis identifies key ca

A:Reference number: S77652; MUID:96059637

A:Accession: S77658

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-659 &lt;PSI&gt;

A:Cross-references: EMBL:246257; NID:g559905; PIDN:CAA86362.1; PID:g559911

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1994

## Query Match

Best Local Similarity 51.4%; Score 36.5; DB 2; Length 659;

Matches 6; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

Qy 1 RPKPQQWF-WLM 11

:|:|:| |||

Db 464 QPGPREWLTWLM 475

## RESULT 105

DNBBEG

DNA-binding protein - bovine herpesvirus 2 (strain BMV)

C:Species: bovine herpesvirus 2

C:Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 07-Jun-1996

C:Accession: A29242

R:Hammerschmidt, W.; Conraths, F.; Mankertz, J.; Pauli, G.; Ludwig, H.; Buhk, H.J.

Virology 165, 388-405, 1988

A:Title: Conservation of a gene cluster including glycoprotein B in bovine herpesvirus t

A:Reference number: A94381; MUID:88306231

A:Accession: A29242

A:Molecule type: DNA

A:Residues: 1-1186 &lt;HAM&gt;

A:Cross-references: GB:M21628

C:Superfamily: herpesvirus DNA-binding protein

C:Keywords: DNA binding

## Query Match

Best Local Similarity 51.4%; Score 36.5; DB 1; Length 1186;

Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 2 PKPQQWF 9

| | |||

Db 834 PNP-QWFW 840

## RESULT 106

S23306

substance P - Atlantic cod

C:Species: Gadus morhua (Atlantic cod)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 18-Aug-2000

C:Accession: S23306

R:Jensen, J.; Conlon, J.M.

Eur. J. Biochem. 206, 659-664, 1992

A:Title: Substance-P-related and neurokinin-A-related peptides from the brain of the cod

A:Reference number: S23186; MUID:92298992

A:Accession: S23306

A:Molecule type: protein

A:Residues: 1-11 &lt;JEN&gt;

A:Experimental source: brain

C:Function:

A:Description: may play a physiological role in the regulation of cardiovascular and  
A:Note: substance P is derived by post-translational processing of preprotachykinin A  
C:Superfamily: unassigned animal peptides  
C:Keywords: neuropeptide; amidated carboxyl end; tachykinin  
F:11/Modified site: amidated carboxyl end (Met) #status predicted

Query Match 50.7%; Score 36; DB 2; Length 11;

Best Local Similarity 54.5%; Pred. No. 6.7;

Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11

:|:|:| |||

Db 1 KRPQQFTGLM 11

## RESULT 107

S45489

H+-transporting ATP synthase (EC 3.6.1.34) protein 8 - European seabass mitochondrion

C:Species: mitochondrion Dicentrarchus labrax (European seabass)

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 07-Dec-1999

C:Accession: S45489

R:Venanzetti, F.; Cecconi, F.; Giorgi, M.; Cesaroni, D.; Sbordoni, V.; Mariottini, P.

Curr. Genet. 26, 139-145, 1994

A:Title: Cloning and characterization of the European seabass, Dicentrarchus labrax,

A:Reference number: S45489; MUID:95094310

A:Accession: S45489

A:Molecule type: DNA

A:Residues: 1-55 &lt;VEN&gt;

A:Cross-references: EMBL:X74147; NID:g521076; PIDN:CAA52244.1; PID:g521077

C:Genetics:

A:Genome: mitochondrion

A:Genetic code: SSCI

C:Superfamily: H+-transporting ATP synthase protein 8

C:Keywords: ATP biosynthesis; hydrolase; membrane-associated complex: mitochondrion;

## Query Match

Best Local Similarity 50.7%; Score 36; DB 2; Length 55;

Matches 4; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 9

:| |||

Db 44 KPSEHSWF 52

## RESULT 108

S77270

hypothetical protein slr0881 - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000

C:Accession: S77270

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys

S.

A:Reference number: S74322; MUID:97061201

A:Accession: S77270

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-103 &lt;KAN&gt;

A:Cross-references: EMBL:D90907; GB:AB001339; NID:g1652618; PIDN:BAAL7604.1; PID:g165

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

C:Superfamily: Synechocystis hypothetical protein slr0881

## Query Match

Best Local Similarity 50.7%; Score 36; DB 2; Length 103;

Matches 3; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 KPQQWF 9

Db 84 RPEHWY 90  
:|:|:|  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 24-Nov-1999  
C:Accession: S43177  
R:Bringaud, F.; Freedland, S.; Liu, X.; Peris, M.; Turck, C.; Simpson, L.  
submitted to the EMBL Data Library, March 1994  
A:Description: Identification of several proteins in mitochondrial nucleoprotein T-co  
A:Reference number: S43177  
A:Accession: S43177  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-187 <BR>  
A:Cross-references: EMBL:Z31697; NID:g469151; PID:g469152  
C:Superfamily: Leishmania tarentolae p18 protein

Query Match 50.7%; Score 36; DB 2; Length 187;  
Best Local Similarity 36.4%; Pred. No. 1.1e+02;  
Matches 4; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
:|:|:|  
Db 113 KPNEESWTWVM 123  
:|:|:|

RESULT 112  
Tl6544  
hypothetical protein K03C7.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Mar-2000  
A:Accession: Tl6544  
R:Leimbach, D.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid K03C7.  
A:Reference number: Z18532  
A:Accession: Tl6544  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-268 <LEI>  
A:Cross-references: EMBL:U40059; NID:g1055170; PID:g1055172; PICN:AAA81139.1; CESP:K0  
C:Superfamily: unassigned fork head proteins; fork head DNA-binding domain homology

Query Match 50.7%; Score 36; DB 2; Length 268;  
Best Local Similarity 55.6%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 9  
:|:|:|  
Db 116 RHRPDQGW 124  
:|:|:|

RESULT 113  
A82185  
glycerol-3-phosphate ABC transporter, permease protein VC1551 [imported] - Vibrio cho  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
A:Accession: A82185  
R:Heidelberger, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.  
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers  
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82095; MUID:20406833  
A:Accession: A82185  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-280 <HEI>  
A:Cross-references: GB:AE004233; GB:AE003852; NID:g9656055; PIDN:AAF94705.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC1551

Db 84 RPEHWY 90  
:|:|:|  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 28-Jul-2000  
C:Accession: C75435  
R:White, O.; Eisen, J.A.; Heidelberger, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: C75435  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-154 <WHI>  
A:Cross-references: GB:AE001961; GB:AE000513; NID:g6458843; PIDN:AAF10689.1; PID:g645885  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1108  
A:Map position: 1  
C:Superfamily: Deinococcus radiodurans hypothetical protein DR1108

Query Match 50.7%; Score 36; DB 2; Length 154;  
Best Local Similarity 54.5%; Pred. No. 91;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPOQWFWM 11  
:|:|:|  
Db 46 RQOPQAFWLL 56  
:|:|:|

RESULT 110  
E83140  
phosphatidylglycerophosphatase A PA4050 [imported] - Pseudomonas aeruginosa (strain PA01  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
A:Accession: E83140  
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337  
A:Accession: E83140  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-171 <STO>  
A:Cross-references: GB:AE004821; GB:AE004091; NID:g9950236; PIDN:AAG07437.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: pg; PA4050  
C:Superfamily: conserved hypothetical protein H11306

Query Match 50.7%; Score 36; DB 2; Length 171;  
Best Local Similarity 50.0%; Pred. No. 1e+02;  
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQWFWM 11  
:|:|:|  
Db 107 PEGWWLL 114  
:|:|:|

RESULT 111  
S43177  
p18 protein - Leishmania tarentolae  
C:Species: Leishmania tarentolae



A;Map position: 1

Query Match 50.7%; Score 36; DB 2; Length 280;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 PKPQOWFLM 11  
| | | | |  
Db 105 PYASAWFLI 114

RESULT 114

A75511 conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: A75511  
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896  
A:Accession: A75511  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-287 <WHI>  
A:Cross-references: GB:AE001909; GB:AE000513; NID:96458188; PIDN:AAF10080.1; PID:9645819  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0500  
A:Map position: 1  
C:Superfamily: Mycoplasma hypothetical protein MG326

Query Match 50.7%; Score 36; DB 2; Length 287;  
Best Local Similarity 71.4%; Pred. No. 1.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQOW 7  
| | | | |  
Db 64 QPSPOQW 70

RESULT 115

I64053 membrane-bound lytic transglycosylase homolog - Haemophilus influenzae (strain Rd KW20)  
C:Species: Haemophilus influenzae  
C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999  
C:Accession: I64053  
R;Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
Science 269, 496-512, 1995  
A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,  
A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
A:Reference number: A64000; MUID:95350630  
A:Accession: I64053  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-318 <TIGR>  
A:Cross-references: GB:U32705; GB:I42023; NID:gi573156; PIDN:AAC21868.1; PID:gi573159; T

Query Match 50.7%; Score 36; DB 2; Length 318;  
Best Local Similarity 40.0%; Pred. No. 1.9e+02;  
Matches 4; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 2 PKPQOWFLM 11  
| | | | |  
Db 295 PAPEQYVIL 304

RESULT 116

S76408 hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76408  
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima,  
O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76408  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-330 <KAN>  
A:Cross-references: EMBL:D90915; GB:AB001339; NID:gi1653604; PIDN:BAAL8537.1; PID:dl01  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 50.7%; Score 36; DB 2; Length 330;  
Best Local Similarity 71.4%; Pred. No. 1.9e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQOW 7  
| | | | |  
Db 64 RPQPNW 70

RESULT 117

S44995 pectate lyase - Erwinia carotovora  
C:Species: Erwinia carotovora  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 08-Oct-1999  
C:Accession: S44995  
R;Heikkinen, R.; Flego, D.; Pirhonen, M.; Karlsson, M.B.; Eriksson, A.; Mae, A.; KO  
submitted to the EMBL Data Library, May 1994  
A:Description: Characterization of a novel pectate lyase from Erwinia carotovora subs  
A:Reference number: S44995  
A:Accession: S44995  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-347 <HEI>  
A:Cross-references: EMBL:X79232; NID:9488382; PIDN:CAA55814.1; PID:9488383

Query Match 50.7%; Score 36; DB 2; Length 347;  
Best Local Similarity 44.4%; Pred. No. 2e+02;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 2 PKPQOWFL 10  
| | | | |  
Db 82 PKSDYVYV 90

RESULT 118

I59347 hypothetical glutathione transporter, sinusoidal - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 12-Mar-1999 #sequence\_revision 12-Mar-1999 #text\_change 31-Dec-2000  
C:Accession: I59347  
R;Yi, J.R.; Lu, S.; Fernandez-Checa, J.; Kaplowitz, N.  
Proc. Natl. Acad. Sci. U.S.A. 92, 1495-1499, 1995  
A:Title: Expression cloning of the cDNA for a polypeptide associated with rat hepatic  
A:Reference number: I59347; MUID:95183492  
A:Accession: I59347  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-353 <YIJ>  
A:Cross-references: EMBL:U16358; NID:9687585; PIDN:AAA62498.1; PID:9687586  
C:Comment: This sequence is very similar to two adjoining hypothetical sequences from

C;Genetics:  
A;Gene: Ragsht

Query Match 50.7%; Score 36; DB 4; Length 353;  
Best Local Similarity 50.0%; Pred. No. 2.1e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
  
QY 2 PKPOQWFWM 11  
| | | | | | | | | |  
DB 124 PVPMTWFMM 133  
  
RESULT 119  
S34173  
Wnt-5c protein - African clawed frog  
C;Species: Xenopus laevis (African clawed frog)  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 16-Jul-1999  
C;Accession: S34173; S45242  
R;Koster, J.G.; Kuiken, G.A.; Stegeman, B.; Peterson, J.; Eizema, K.; Stabel, L.; Dekker  
submitted to the EMBL Data Library, June 1993  
A;Description: Differential Xwt-5c expression during early development of Xenopus laevis  
A;Reference number: S34173  
A;Accession: S34173  
A;Molecule type: mRNA  
A;Residues: 1-360 <ROS>  
A;Cross-references: EMBL:X73510; NID:g313267; PIDN:CAA51916.1; PID:g313268  
R;Kuiken, G.A.; Bertens, P.J.A.; Peterson-Maduro, J.; Veenstra, G.J.C.; Koster, J.G.; De  
Nucleic Acids Res. 22, 1675-1680, 1994  
A;Title: The promoter of the Xwt-5c gene contains octamer and AP-2 motifs functional in  
A;Reference number: S45242; MUID:94261437  
A;Accession: S45242  
A;Molecule type: DNA  
A;Residues: 1-28 <KUI>  
C;Superfamily: Int-1 transforming protein

Query Match 50.7%; Score 36; DB 2; Length 360;  
Best Local Similarity 54.5%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;  
  
QY 1 RPK--PQQWFW 9  
| | | | | | | | | |  
DB 149 RPKDLPRDWLW 159

RESULT 120  
A48914  
proto-oncogene Wnt-5A precursor - human  
C;Species: Homo sapiens (man)  
C;Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 16-Jul-1999  
C;Accession: A48914  
R;Clark, C.C.; Cohen, I.; Eichstetter, I.; Cannizzaro, L.A.; McPherson, J.D.; Wasmuth, J.  
Genomics 18, 249-260, 1993  
A;Title: Molecular cloning of the human proto-oncogene Wnt-5A and mapping of the gene (W  
A;Reference number: A48914; MUID:94116991  
A;Accession: A48914  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-365 <CLA>  
A;Cross-references: GB:L20861; NID:g348917; PIDN:AAA16842.1; PID:g348918  
C;Genetics:  
A;Gene: GDB:WNT5A  
A;Cross-references: GDB:141726; OMIM:164975  
A;Map position: 3p21-3p14  
C;Superfamily: Int-1 transforming protein

Query Match 50.7%; Score 36; DB 2; Length 365;  
Best Local Similarity 54.5%; Pred. No. 2.1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;  
  
QY 1 RPK--PQQWFW 9

DB 154 RPKDLPRDWLW 164  
| | | | | | | | | |

RESULT 121  
S75038  
hypothetical protein sll1611 - Synechocystis sp. (strain PCC 6803)  
C;Species: Synechocystis sp.  
A;Variety: PCC 6803  
C;Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C;Accession: S75038  
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
S.  
A;Reference number: S74322; MUID:97061201  
A;Accession: S75038  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-369 <KAN>  
A;Cross-references: EMBL:D90910; GB:AB001339; NID:g1652956; PIDN:BAA17900.1; PID:d101  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C;Genetics:  
A;Start codon: GTG

Query Match 50.7%; Score 36; DB 2; Length 369;  
Best Local Similarity 71.4%; Pred. No. 2.2e+02;  
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 4 PQQWFWL 10  
| | | | | | | | | |  
DB 52 PQQWLWL 58

RESULT 122  
E36470  
Wnt-5b protein - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 19-Apr-1991 #sequence\_revision 19-Apr-1991 #text\_change 16-Jul-1999  
C;Accession: E36470  
R;Gavin, B.J.; McMahon, J.A.; McMahon, A.P.  
Genes Dev. 4, 2319-2332, 1990  
A;Title: Expression of multiple novel Wnt-1/int-1-related genes during fetal and adul  
A;Reference number: A36470; MUID:91122634  
A;Accession: E36470  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-372 <GAV>  
A;Cross-references: GB:M89799; NID:g202405; PIDN:AAA40568.1; PID:g202406  
C;Superfamily: int-1 transforming protein

Query Match 50.7%; Score 36; DB 2; Length 372;  
Best Local Similarity 54.5%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;  
  
QY 1 RPK--PQQWFW 9  
| | | | | | | | | |  
DB 161 RPKDLPRDWLW 171

RESULT 123  
D36470  
Wnt-5a protein - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 19-Apr-1991 #sequence\_revision 19-Apr-1991 #text\_change 16-Jul-1999  
C;Accession: D36470  
R;Gavin, B.J.; McMahon, J.A.; McMahon, A.P.  
Genes Dev. 4, 2319-2332, 1990  
A;Title: Expression of multiple novel Wnt-1/int-1-related genes during fetal and adul  
A;Reference number: A36470; MUID:91122634

A:Accession: D36470  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-379 <GAV>  
A:Cross-references: GB:M89798; NID:g202403; PIDN:AAA04567.1; PID:g202404  
C:Superfamily: int-1 transforming protein

Query Match 50.7%; Score 36; DB 2; Length 379;  
Best Local Similarity 54.5%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Qy 1 RPK--PQQWF 9  
||| | : |  
Db 168 RPKDLPDRLW 178

## RESULT 124

A71390  
ubiquinol--cytochrome-c reductase (EC 1.10.2.2) cytochrome b - common lancelet mitochondrion  
C:Species: mitochondrion Branchiostoma lanceolatum (common lancelet)  
C:Date: 03-Aug-1998 #sequence\_revision 03-Aug-1998 #text\_change 20-Jun-2000  
C:Accession: A71390  
R:Spruyt, N.; Delarbre, C.; Gachelin, G.; Laudet, V.  
Nucleic Acids Res. 26, 3279-3285, 1998  
A:Title: Complete sequence of the amphioxus (Branchiostoma lanceolatum) mitochondrial gene  
A:Reference number: A71390; MUID:98292550

A:Accession: A71390

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-380 <SPR>

A:Cross-references: GB:Y16474; NID:g3292989; PIDN:CAA76246.1; PID:g3292990

C:Genetics:

A:Gene: cytb

A:Genome: mitochondrion

A:Genetic code: SGC4

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol  
C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion;  
F:12-210/Domain: cytochrome b6 homology <CB6>

F:37-53/Domain: transmembrane #status predicted <TM1>

F:82-100/Domain: transmembrane #status predicted <TM2>

F:118-134/Domain: transmembrane #status predicted <TM3>

F:179-201/Domain: transmembrane #status predicted <TM4>

F:222-340/Domain: plastoquinol--plastocyanin reductase 17K protein homology <17K>

F:230-246/Domain: transmembrane #status predicted <TM5>

F:289-305/Domain: transmembrane #status predicted <TM6>

F:324-344/Domain: transmembrane #status predicted <TM7>

F:354-371/Domain: transmembrane #status predicted <TM8>

F:84,183/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:98,197/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 50.7%; Score 36; DB 2; Length 380;  
Best Local Similarity 63.6%; Pred. No. 2.2e+02;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 11  
||| | ||||  
Db 319 RPLAQVLF 329

## RESULT 125

S70594  
ubiquinol--cytochrome-c reductase (EC 1.10.2.2) cytochrome b - starfish (Asterina pectinifera)  
C:Species: mitochondrion Asterina pectinifera  
C:Date: 14-Feb-1997 #sequence\_revision 13-Mar-1997 #text\_change 04-Mar-2000  
C:Accession: S70594

R:Asakawa, S.; Himeno, H.; Miura, K.; Watanabe, K.

Genetics 140, 1047-1060, 1995

A:Title: Nucleotide sequence and gene organization of the starfish *Asterina pectinifera*

A:Reference number: S70589; MUID:95402698

A:Accession: S70594

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-380 <ASA>

A:Cross-references: EMBL:D16387

A:Experimental source: ovary

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1993

C:Genetics:

A:Gene: cytb

A:Genome: mitochondrion

A:Genetic code: SGC8

A:Start codon: ATT

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion

F:13-341/Domain: cytochrome b homology <CBH>

F:13-211/Domain: cytochrome b6 homology <CB6>

F:38-54/Domain: transmembrane #status predicted <TM1>

F:83-101/Domain: transmembrane #status predicted <TM2>

F:119-135/Domain: transmembrane #status predicted <TM3>

F:180-202/Domain: transmembrane #status predicted <TM4>

F:223-341/Domain: plastoquinol--plastocyanin reductase 17K protein homology <17K>

F:231-247/Domain: transmembrane #status predicted <TM5>

F:290-306/Domain: transmembrane #status predicted <TM6>

F:325-345/Domain: transmembrane #status predicted <TM7>

F:355-371/Domain: transmembrane #status predicted <TM8>

F:85,184/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:99,198/Binding site: heme iron (His) (axial ligands) (high potential) #status predicted

Query Match 50.7%; Score 36; DB 2; Length 380;

Best Local Similarity 54.5%; Pred. No. 2.2e+02;

Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 RPKPQQWF 11

||| | |||

Db 320 RPLAQVLF 330

## RESULT 126

S59093  
ubiquinol--cytochrome-c reductase (EC 1.10.2.2) cytochrome b - red alga (Chondrus cri

N:Alternate names: apocytochrome b

C:Species: mitochondrion Chondrus crispus (carrageen)

C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 20-Jun-2000

C:Accession: S59093

R:Leblanc, C.; Boyen, C.; Richard, O.; Bonnard, G.; Grienemberger, J.M.; Kloareg, B.

J. Mol. Biol. 250, 484-495, 1995

A:Title: Complete sequence of the mitochondrial DNA of the rhodophyte *Chondrus crispus*

A:Reference number: S59078; MUID:95341681

A:Accession: S59093

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-381 <LEB>

A:Cross-references: EMBL:Z47547; NID:g1019057; PIDN:CAA87609.1; PID:g1334487

A:Experimental source: female gametophytes

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1995

C:Genetics:

A:Gene: cob

A:Genome: mitochondrion

A:Genetic code: SGC3

C:Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol

C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; mitochondrion

F:9-339/Domain: cytochrome b homology <CBH>

F:9-209/Domain: cytochrome b6 homology <CB6>

F:34-50/Domain: transmembrane #status predicted <TM1>

F:79-97/Domain: transmembrane #status predicted <TM2>

F:117-133/Domain: transmembrane #status predicted <TM3>

F:178-200/Domain: transmembrane #status predicted <TM4>

F:221-339/Domain: plastoquinol--plastocyanin reductase 17K protein homology <17K>

F:229-245/Domain: transmembrane #status predicted <TM5>

F:288-304/Domain: transmembrane #status predicted <TM6>

F:323-343/Domain: transmembrane #status predicted <TM7>

F:353-369/Domain: transmembrane #status predicted <TM8>

F:81,182/Binding site: heme iron (His) (axial ligands) (low potential) #status predicted

F:95.196/Binding site: heme iron (His) (axial ligands) (high potential) #status,predicted

Query Match 50.7%; Score 36; DB 2; Length 381;  
Best Local Similarity 44.4%; Pred. No. 2.2e+02;  
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 KPQOWFWLM 11  
| : | : | :  
Db 107 KPRHWVVI 115

RESULT 127  
T39028  
citrate synthase precursor, mitochondrial - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T39028  
R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, February 1996  
A:Reference number: Z21750  
A:Accession: T39028  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-473 <DEV>  
A:Cross-references: EMBL:Z69731; PIDN:CAA93617.2; GSPDB:GN00066; SPDB:SPAC6C3.04  
A:Experimental source: strain 972h-; cosmid c6C3  
C:Genetics:  
A:Gene: SPDB:SPAC6C3.04  
A:Map position: 1  
A:Genome: nuclear  
C:Superfamily: citrate (sl)-synthase  
C:Keywords: mitochondrion

Query Match 50.7%; Score 36; DB 2; Length 473;  
Best Local Similarity 45.5%; Pred. No. 2.8e+02;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPOWFWLM 11  
| : | : | : | :  
Db 122 QPLPESLFWLL 132

RESULT 128  
JC4532  
cytochrome P450 4F4 protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 15-Feb-1996 #sequence\_revision 19-Apr-1996 #text\_change 28-Jul-2000  
C:Accession: JC4532  
R:Kawashima, H.; Strobel, H.W.  
Biochem. Biophys. Res. Commun. 217, 1137-1144, 1995  
A:Title: cDNA cloning of three new forms of rat brain cytochrome P450 belonging to the C  
A:Reference number: JC4532; MUID:96125358  
A:Accession: JC4532  
A:Molecule type: mRNA  
A:Residues: 1-522 <KAW>  
A:Cross-references: GB:U39206; NID:g1146435; PIDN:AAC52358.1; PID:g1146436  
A:Experimental source: brain  
C:Superfamily: human cytochrome P450 CYP4B1; cytochrome P450 homology  
C:Keywords: brain; chromoprotein; heme; iron; metalloprotein  
F:325-490/Domain: cytochrome P450 homology <P45>  
F:468/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 50.7%; Score 36; DB 2; Length 522;  
Best Local Similarity 40.0%; Pred. No. 3.1e+02;  
Matches 4; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPOWFWLM 11  
| : | : | : | :  
Db 52 PQPKWVFL 61

RESULT 129

T35047  
hypothetical protein SC4G2.12c SC4G2.12c - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 01-Dec-2000  
C:Accession: T35047; T30206  
R:Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1998  
A:Reference number: Z21566  
A:Accession: T35047  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-605 <SEE>  
A:Cross-references: EMBL:AL031371; PIDN:CAA20549.1; GSPDB:GN00070; SCOEDB:SC4G2.12c  
A:Experimental source: strain A3(2)  
R:Bedford, D.J.; Laity, C.; Buttner, M.J.  
J. Bacteriol. 177, 4681-4689, 1995  
A:Title: Two genes involved in the phase-variable phi C31 resistance mechanism of Str  
A:Reference number: Z20777; MUID:95370146  
A:Accession: T30206  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 304-605 <BED>  
A:Cross-references: EMBL:L37531; NID:g576537; PID:g576541; PIDN:AAB00368.1  
C:Genetics:  
A:Gene: SCOEDB:SC4G2.12c

Query Match 50.7%; Score 36; DB 2; Length 605;  
Best Local Similarity 57.1%; Pred. No. 3.5e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 KPQOWFW 9  
| : | : | : | :  
Db 396 KTEQWY 402

RESULT 130  
E69334  
acetyl-CoA synthetase (acs-3) homolog - Archaeoglobus fulgidus  
C:Species: Archaeoglobus fulgidus  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: E69334  
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod  
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E  
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
Nature 390, 364-370, 1997  
A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes,  
Smith, H.O.; Woese, C.R.; Venter, J.C.  
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
A:Reference number: A69250; MUID:98049343  
A:Accession: E69334  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-643 <KLE>  
A:Cross-references: GB:AE001058; GB:AE000782; NID:g2689381; PIDN:AAB90564.1; PID:g264  
C:Superfamily: acetate--CoA ligase; acetate--CoA ligase homology  
F:108-609/Domain: acetate--CoA ligase homology <ACL>

Query Match 50.7%; Score 36; DB 2; Length 643;  
Best Local Similarity 57.1%; Pred. No. 3.8e+02;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQOWFWL 10  
| : | : | : | :  
Db 377 PEVWYWL 383

RESULT 131  
T01573  
earl protein - maize

C:Species: Zea mays (maize)  
C:Date: 19-Feb-1999 #sequence\_revision 19-Feb-1999 #text\_change 29-Oct-1999  
C:Accession: T01573  
R:Veit, B.E.; Briggs, S.P.; Schmidt, R.J.; Yanofsky, M.F.; Hake, S.  
Nature 393, 166-169, 1998  
A:Title: Regulation of leaf initiation by the terminal earl gene of maize.  
A:Reference number: Z14351; MUID:98264681  
A:Accession: T01573  
A:Status: preliminary; translated from GB/EMBL/DDBT  
A:Molecule type: mRNA  
A:Residues: 1-656 <VEI>  
A:Cross-references: EMBL:AF047852; NID:g3153236; PIDN:AAC39463.1; PID:g3153237  
A:Experimental source: cultivar B73  
C:Genetics:  
A:Map position: 3

Query Match 50.7%; Score 36; DB 2; Length 656;  
Best Local Similarity 62.5%; Pred. No. 3.8e+02;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 PKPQQWF 9  
| | | | |  
DB 172 PTPQAMDW 179

RESULT 132  
A75542  
conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
C:Accession: A75542  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Sheth, H.O.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75520; MUID:20036896  
A:Accession: A75542  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-670 <WHI>  
A:Cross-references: GB:AE001886; GB:AE000513; NID:g6457921; PIDN:AAF09837.1; PID:g645792  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0250  
A:Map position: 1

Query Match 50.7%; Score 36; DB 2; Length 670;  
Best Local Similarity 66.7%; Pred. No. 3.9e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 QQWF 10  
| | | | |  
DB 95 QRWF 100

RESULT 133  
H82381  
toxin secretion ATP-binding protein VCA1084 [imported] - Vibrio cholerae (strain N16961)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: H82381  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Richardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.  
L. R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: A82035; MUID:20406833  
A:Accession: H82381  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-704 <HEI>  
A:Cross-references: GB:AE004433; GB:AE003853; NID:g9658519; PIDN:AAF96977.1; GSPDB:GN  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VCA1084  
A:Map position: 2

Query Match 50.7%; Score 36; DB 2; Length 704;  
Best Local Similarity 44.4%; Pred. No. 4.1e+02;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
| | | | |  
DB 136 KPRDGHWF 144

RESULT 134  
C85547  
probable cytoplasmic membrane export protein Z0634 [imported] - Escherichia coli (str  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
C:Accession: C85547  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: C85547  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-720 <SPO>  
A:Cross-references: GB:AE005174; NID:gl2513369; PIDN:AAG54839.1; GSPDB:GN00145; UWGP:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z0634

Query Match 50.7%; Score 36; DB 2; Length 720;  
Best Local Similarity 55.6%; Pred. No. 4.2e+02;  
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQQWF 9  
| | | | |  
DB 141 RPYQANWF 149

RESULT 135  
A70010  
NADH dehydrogenase homolog yufT - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 20-Jun-2000  
C:Accession: A70010  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M  
Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino  
A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl  
A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida  
A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili  
A:Reference number: A69580; MUID:98044033  
A:Accession: A70010  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-774 <KUN>

A:Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15149.1; PID:g2635656  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: yufl  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5

Query Match 50.7%; Score 36; DB 2; Length 774;  
Best Local Similarity 66.7%; Pred. No. 4.5e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 QWFWLM 11  
:||||:  
DB 245 QWFWIV 250

RESULT 136  
S61395  
probable Na<sup>+</sup>/H<sup>+</sup>-exchanging protein system component - Bacillus sp. (strain C-125)  
N:Alternate names: NADH dehydrogenase (ubiquinone) chain 5 homolog; probable Na<sup>+</sup>/H<sup>+</sup> antiporter  
C:Species: Bacillus sp.  
A:Variety: strain C-125  
C:Date: 01-Mar-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Jun-2000  
C:Accession: S61395; S61392  
R:Hamamoto, T.; Hashimoto, M.; Hino, M.; Kitada, M.; Seto, Y.; Kudo, T.; Horikoshi, K.  
submitted to the EMBL Data Library, June 1994  
A:Description: Characterization of a gene responsible for Na<sup>+</sup>/H<sup>+</sup> antiporter system of al  
A:Reference number: S61395  
A:Accession: S61395  
A:Molecule type: DNA  
A:Residues: 1-804 <HAM>  
A:Cross-references: EMBL:D31823; NID:g854654; PIDN:BAA06609.1; PID:g854655  
A:Experimental source: Bacillus sp. strain C-125  
R:Hamamoto, T.; Hashimoto, M.; Hino, M.; Kitada, M.; Seto, Y.; Kudo, T.; Horikoshi, K.  
Mol. Microbiol. 14, 939-946, 1994  
A:Title: Characterization of a gene responsible for the Na<sup>+</sup>/H<sup>+</sup> antiporter system of  
A:Reference number: S61392; MUID:95231300  
A:Accession: S61392  
A:Molecule type: DNA  
A:Residues: 123-166; 214-267; 282-314; 355-474 <HAM>  
A:Cross-references: EMBL:D31823  
A:Experimental source: Bacillus sp. strain C-125  
A:Note: only a part of the coding sequence is given  
C:Genetics:  
A:Start codon: TTG  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5

Query Match 50.7%; Score 36; DB 2; Length 804;  
Best Local Similarity 66.7%; Pred. No. 4.7e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 QWFWLM 11  
:||||:  
DB 272 EWFLL 277

RESULT 137  
G83814  
Na<sup>+</sup>/H<sup>+</sup> antiporter BH1319 [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: G83814  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: AB3650; MUID:20263314  
A:Accession: G83814  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-804 <STO>  
A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA805038.1; GSPDB:GN00  
A:Experimental source: strain C-125

C:Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15149.1; PID:g2635656  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: BH1319  
C:Superfamily: NADH dehydrogenase (ubiquinone) chain 5

Query Match 50.7%; Score 36; DB 2; Length 804;  
Best Local Similarity 66.7%; Pred. No. 4.7e+02;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 QWFWLM 11  
:||||:  
DB 272 EWFLL 277

RESULT 138  
S75776  
pled protein - Synechocystis sp. (strain PCC 6803)  
N:Alternate names: protein slr0829  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S75776  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S75776  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-829 <KAN>  
A:Cross-references: EMBL:D64003; GB:AB001339; NID:g1001200; PIDN:BAA10511.1; PID:d101  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: pld  
A:Start codon: GTG

Query Match 50.7%; Score 36; DB 2; Length 829;  
Best Local Similarity 54.5%; Pred. No. 4.8e+02;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 RPKPQWFWLM 11  
:||||:  
DB 325 RDARQQRWLL 335

RESULT 139  
S77086  
hypothetical protein sl10737 - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S77086  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S77086  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-861 <KAN>  
A:Cross-references: EMBL:D64005; GB:AB001339; NID:g1001779; PIDN:BAA10778.1; PID:g100  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Superfamily: Synechocystis hypothetical protein sl10737

Query Match 50.7%; Score 36; DB 2; Length 861;  
Best Local Similarity 83.3%; Pred. No. 5e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 QQWFNL 10  
I :|||I  
Db 57 QTWFNL 62

## RESULT 140

T14761

hypothetical protein DKFZp434K233.1 - human (fragment)

C:Species: Homo sapiens (man)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T14761  
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, August 1999  
A:Reference number: Z18181  
A:Accession: r14761  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-959 <WAM>  
A:Cross-references: EMBL:AL110222  
A:Experimental source: adult testis; clone DKFZp434K233  
C:Genetics:  
A:Note: DKFZp434K233.1

Query Match 50.7%; Score 36; DB 2; Length 959;  
Best Local Similarity 62.5%; Pred. No. 5.6e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKQQWF 8  
:|||I  
Db 682 KPKPPAWF 689

## RESULT 141

T25033

hypothetical protein T20D3.9 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
C:Accession: T25033  
R:Lloyd, C.  
submitted to the EMBL Data Library, December 1995  
A:Reference number: Z19971  
A:Accession: T25033  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-1038 <WIL>  
A:Cross-references: EMBL:Z68220; PIDN:CAA92492.2; GSPDB:GN000022; CESP:T20D3.9  
A:Experimental source: clone T20D3  
C:Genetics:  
A:Gene: CESP:T20D3.9  
A:Map position: 4  
A:Introns: 36/3; 341/3; 380/1; 574/2; 771/2; 966/1; 1010/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein T20D3.9

Query Match 50.7%; Score 36; DB 2; Length 1038;  
Best Local Similarity 44.4%; Pred. No. 6e+02;  
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 RPKQQWFNL 10  
I :|||I  
Db 246 PRPSKFFW 254

## RESULT 142

T22982

hypothetical protein F59B10.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Sep-2000  
C:Accession: T22982; T27953  
R:Lloyd, C.  
submitted to the EMBL Data Library, March 1995  
A:Reference number: Z19646

A:Accession: T22982  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-1039 <WIL>  
A:Cross-references: EMBL:Z48716; PIDN:CAA88602.2; GSPDB:GN000020; CESP:F59B10.1  
A:Experimental source: clone F59B10  
R:Wilkinson, J.  
submitted to the EMBL Data Library, April 1995  
A:Reference number: Z20445  
A:Accession: T27953  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-1039 <WIL>  
A:Cross-references: EMBL:Z49132; PIDN:CAA88990.2; GSPDB:GN000020; CESP:F59B10.1  
A:Experimental source: clone ZK666  
C:Genetics:  
A:Gene: CESP:F59B10.1  
A:Map position: 2  
A:Introns: 9/1; 73/3; 293/3; 711/1; 754/3; 837/2; 877/3; 927/2  
C:Superfamily: Caenorhabditis elegans hypothetical protein F59B10.1

Query Match 50.7%; Score 36; DB 2; Length 1039;  
Best Local Similarity 54.5%; Pred. No. 6e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 RPKQQWF 11  
:|||I  
Db 387 RNPQKFFEL 397

## RESULT 143

T19214

UDP-glucose--glycoprotein glucosyltransferase (EC 2.4.1.-) precursor F26H9.8 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T19214; T21444  
R:Barlow, K.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19091  
A:Accession: T19214  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-1377 <WIL>  
A:Cross-references: EMBL:Z81467; PIDN:CAB03874.1; GSPDB:GN000019; CESP:F26H9.8  
A:Experimental source: clone C12C8  
R:Baynes, C.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19422  
A:Accession: T21444  
A:Status: preliminary; translated from GB/EMBL/DBDJ  
A:Molecule type: DNA  
A:Residues: 1-1377 <WIL>  
A:Cross-references: EMBL:Z81516; PIDN:CAB04207.1; GSPDB:GN000019; CESP:F26H9.8  
A:Experimental source: clone F26H9

C:Genetics:  
A:Gene: CESP:F26H9.8  
A:Map position: 1  
A:Introns: 40/2; 70/1; 152/2; 318/2; 429/3; 494/2; 538/3; 564/3; 625/3; 654/3; 782/3;  
C:Keywords: glycoprotein; glycosyltransferase; hexosyltransferase

Query Match 50.7%; Score 36; DB 2; Length 1377;  
Best Local Similarity 66.7%; Pred. No. 8e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPKQQWF 9  
:|||I  
Db 1305 PQEWLW 1310

## RESULT 144

T16404

hypothetical protein F48E3.3 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
A;Accession: T16404  
R;Pauley, A.  
submitted to the EMBL Data Library, June 1995  
A;Description: The sequence of C. elegans cosmid F48E3.  
A;Reference number: Z18508  
A;Accession: T16404  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: DNA  
A;Residues: 1-1493 <PAU>  
A;Cross-references: EMBL:U28735; NID:g860712; PIDN:AAA68266.1; CESP:F48E3.3  
A;Experimental source: strain Bristol N2  
C;Genetics:  
A;Gene: CESP:F48E3.3  
A;Introns: 42/2; 72/1; 115/3; 203/3; 391/3; 421/3; 634/1; 712/3; 1017/2; 1190/2; 1438/3

Query Match 50.7%; Score 36; DB 2; Length 1493;  
Best Local Similarity 66.7%; Pred. No. 8.6e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQFW 9  
||| |  
Db 1415 PQEWLW 1420

RESULT 145  
G96736  
hypothetical protein F3117.13 [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
A;Accession: G96736  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719  
A;Accession: G96736  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1674 <STO>  
A;Cross-references: GB:AE0051173; NID:g6714344; PIDN:AAF26036.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: F3117.13  
A;Map position: 1

Query Match 50.7%; Score 36; DB 2; Length 1674;  
Best Local Similarity 66.7%; Pred. No. 9.7e+02;  
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PQQFW 9  
||| |  
Db 1575 PQEWLW 1580

RESULT 146  
T13714  
kakapo gene protein - fruit fly (Drosophila melanogaster) (fragment)  
C;Species: Drosophila melanogaster  
C;Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000  
A;Accession: T13714; T13715  
R;Gregory, S.L.; Brown, N.H.  
J. Cell Biol. 143, 1271-1282, 1998  
A;Title: Kakapo, a gene required for adhesion between and within cell layers in Drosophila

A;Reference number: Z17707; MUID:99054753  
A;Accession: T13714  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 1-2396 <GRE>  
A;Cross-references: EMBL:AJ011924; NID:g3758908; PIDN:CAA09869.1; PID:g3758909  
A;Accession: T13715  
A;Status: preliminary; translated from GB/EMBL/DBDJ  
A;Molecule type: mRNA  
A;Residues: 'KM', 77-79, 'SL', 82, 'E', 'WAKDK', 108-109, 'STILOLD', 116-117, 'DR', 138, 'VLRIA',  
A;Cross-references: EMBL:AJ011925; NID:g3758910; PIDN:CAA09870.1; PID:g3758911  
C;Genetics:  
A;Gene: kak  
A;Cross-references: FlyBase:FBgn0013733  
A;Note: kak

Query Match 50.7%; Score 36; DB 2; Length 2396;  
Best Local Similarity 71.4%; Pred. No. 1.4e+03;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 5 QQWFELM 11  
||| |  
Db 815 QQWAWLL 821

RESULT 147  
S71752  
giant protein p619 - human  
N;Alternate names: chromosome condensation regulator RCC1 homolog p619  
C;Species: Homo sapiens (man)  
C;Date: 29-Jan-1998 #sequence\_revision 06-Feb-1998 #text\_change 21-Jul-2000  
A;Accession: S71752  
R;Rosa, J.L.; Casaroli-Marano, R.P.; Buckler, A.J.; Vilaro, S.; Barbacid, M.  
EMBO J. 15, 4262-4273, 1996  
A;Title: p619, a giant protein related to the chromosome condensation regulator RCC1.  
A;Reference number: S71752; MUID:97015127  
A;Accession: S71752  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-4861 <ROS>  
A;Cross-references: EMBL:U50078; NID:g4220427; PIDN:AAD12586.1; PID:g1477565  
C;Genetics:  
A;Gene: p619  
C;Function:  
A;Description: may play an important role in the regulation of membrane trafficking;  
C;Superfamily: human giant protein p619; ubiquitin--protein ligase homology; WD repeat  
C;Keywords: leucine zipper  
F;1771-1805/Region: leucine zipper motif  
F;3424-3457/Domain: WD repeat homology <WD1>  
F;3743-3776/Domain: WD repeat homology <WD2>  
F;4484-4838/Domain: ubiquitin--protein ligase homology <UBI>

Query Match 50.0%; Score 35.5; DB 2; Length 4861;  
Best Local Similarity 54.5%; Pred. No. 3.3e+03;  
Matches 6; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 2 PKP-QQWFELM 11  
||| |  
Db 1129 PQAQSWWLV 1139

Search completed: April 1, 2002, 16:19:21  
Job time: 112 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 1, 2002, 16:20:17 ; Search time 13.22 Seconds  
(without alignments)  
30.508 Million cell updates/sec

Title: US-09-988-792-2  
Perfect score: 71  
Sequence: 1 RPKPQQWFWM 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues  
Total number of hits satisfying chosen parameters: 92

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 50%  
Maximum Match 100%  
Listing first 1000 summaries

Database : SwissProt\_39.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID         | Description         |
|------------|-------|-------------|--------|------------|---------------------|
| 1          | 48    | 67.6        | 11     | TKNA_HORSE | P01290 equus caball |
| 2          | 48    | 67.6        | 115    | TKN1_RABIT | P41540 oryctolagus  |
| 3          | 48    | 67.6        | 129    | TKN1_HUMAN | P20366 homo sapien  |
| 4          | 48    | 67.6        | 130    | TKN1_BOVIN | P01289 bos taurus   |
| 5          | 48    | 67.6        | 130    | TKN1_MESAU | Q60541 mesocricetu  |
| 6          | 48    | 67.6        | 130    | TKN1_MOUSE | P41539 mus musculu  |
| 7          | 48    | 67.6        | 130    | TKN1_RAT   | P06767 rattus norv  |
| 8          | 45    | 63.4        | 11     | TKNA_CHICK | P19850 gallus gall  |
| 9          | 44    | 62.0        | 436    | CYB_RHOC   | P08502 rhodobacter  |
| 10         | 43    | 60.6        | 902    | VEF_GVHA   | P54232 heliothis a  |
| 11         | 42    | 59.2        | 377    | DCHE_ENTAE | P28577 enterobacte  |
| 12         | 41    | 57.7        | 55     | ATP8_SQUAC | Q92250 squalius aca |
| 13         | 40    | 56.3        | 411    | CIW2_MOUSE | P97438 mus musculu  |
| 14         | 40    | 56.3        | 426    | CIW2_HUMAN | O95059 homo sapien  |
| 15         | 40    | 56.3        | 538    | CIWA_HUMAN | P57789 homo sapien  |
| 16         | 40    | 56.3        | 538    | CIWA_RAT   | O9j1s4 rattus norv  |
| 17         | 40    | 56.3        | 665    | CNG_DRONE  | Q24278 drosophila   |
| 18         | 39    | 54.9        | 361    | FUT3_HUMAN | P21217 homo sapien  |
| 19         | 39    | 54.9        | 372    | FUT3_PANTR | O19058 pan troglod  |
| 20         | 39    | 54.9        | 374    | FUT5_HUMAN | Q11128 homo sapien  |
| 21         | 39    | 54.9        | 374    | FUT5_PANTR | P56433 pan troglod  |
| 22         | 39    | 54.9        | 440    | CYB_PARDE  | P05418 paracoccus   |
| 23         | 39    | 54.9        | 816    | NPA2_MOUSE | P97460 mus musculu  |
| 24         | 39    | 54.9        | 824    | NPA2_HUMAN | O99743 homo sapien  |
| 25         | 39    | 54.9        | 846    | CLOC_HUMAN | O15516 homo sapien  |
| 26         | 39    | 54.9        | 855    | CLOC_MOUSE | O08785 mus musculu  |
| 27         | 39    | 54.9        | 1023   | CLOC_DRONE | O61735 drosophila   |
| 28         | 38    | 53.5        | 323    | MSBB_ECOLI | P24205 escherichia  |
| 29         | 38    | 53.5        | 330    | LSPI_MOUSE | P19973 mus musculu  |
| 30         | 38    | 53.5        | 405    | FUT4_HUMAN | P22083 homo sapien  |
| 31         | 38    | 53.5        | 433    | FUT4_MOUSE | Q11127 mus musculu  |
| 32         | 38    | 53.5        | 433    | FUT4_RAT   | Q62994 rattus norv  |
| 33         | 38    | 53.5        | 533    | YADC_SCHPO | Q09837 schizosacch  |

|    |      |      |   |            |                     |
|----|------|------|---|------------|---------------------|
| 34 | 53.5 | 789  | 1 | ARNT_HUMAN | P27540 homo sapien  |
| 35 | 53.5 | 901  | 1 | VEF_GVPU   | P41723 pseudalattia |
| 36 | 53.5 | 901  | 1 | VEF_GVTN   | P29998 trichoplusi  |
| 37 | 52.8 | 298  | 1 | Y812_ARCFU | Q29446 archaeoglob  |
| 38 | 52.8 | 375  | 1 | DNBI_HSV1  | Q03444 equine herp  |
| 39 | 52.8 | 1196 | 1 | DNBI_HSV1  | P04296 herpes simp  |
| 40 | 52.8 | 1196 | 1 | DNBI_HSV1K | P17460 herpes simp  |
| 41 | 52.8 | 1196 | 1 | DNBI_HSV1K | P17470 herpes simp  |
| 42 | 52.8 | 1196 | 1 | DNBI_HSV2H | P89452 herpes simp  |
| 43 | 52.8 | 1137 | 1 | DNBI_HSV2  | P86384 herpes simp  |
| 44 | 52.8 | 1204 | 1 | DNBI_VZVD  | P36384 herpes simp  |
| 45 | 52.8 | 1209 | 1 | DNBI_HSV2B | P28932 equine herp  |
| 46 | 52.8 | 11   | 1 | TKNA_ONCMY | P28932 equine herp  |
| 47 | 52.1 | 11   | 1 | TKNA_ONCMY | P28932 equine herp  |
| 48 | 52.1 | 175  | 1 | HRB3_XANCV | P36558 escherichia  |
| 49 | 52.1 | 253  | 1 | HRB3_XANCV | P80152 xanthomonas  |
| 50 | 52.1 | 342  | 1 | FUT7_HUMAN | Q11130 homo sapien  |
| 51 | 52.1 | 403  | 1 | YCGF_ECOLI | P75990 escherichia  |
| 52 | 52.1 | 404  | 1 | CYB_MARPO  | P26852 marchantia   |
| 53 | 52.1 | 655  | 1 | NEC3_MOUSE | P29121 mus musculu  |
| 54 | 52.1 | 663  | 1 | TRA_BPMU   | P07636 bacterioph   |
| 55 | 52.1 | 681  | 1 | TBR1_MOUSE | Q64336 mus musculu  |
| 56 | 52.1 | 682  | 1 | TBR1_HUMAN | Q16650 homo sapien  |
| 57 | 52.1 | 747  | 1 | AMDI_HUMAN | P23109 homo sapien  |
| 58 | 52.1 | 956  | 1 | METE_SOLSC | Q42662 solenostemo  |
| 59 | 52.1 | 971  | 1 | AREA_GIBFU | Q01168 magnaporthe  |
| 60 | 52.1 | 1002 | 1 | DOR_DRONE  | P78688 gibberella   |
| 61 | 52.1 | 1239 | 1 | DPOG_HUMAN | Q24314 drosophila   |
| 62 | 52.1 | 1629 | 1 | ATP8_HUMAN | P34098 homo sapien  |
| 63 | 52.1 | 1629 | 1 | ATP8_HUMAN | Q9p2n4 homo sapien  |
| 64 | 51.4 | 537  | 1 | CPF6_RAT   | P51871 rattus norv  |
| 65 | 51.4 | 564  | 1 | NOX1_HUMAN | O9y5s8 homo sapien  |
| 66 | 51.4 | 659  | 1 | Y102_MYCLE | P53525 mycobacteri  |
| 67 | 51.4 | 1186 | 1 | DNBI_HSVB2 | P12639 bovine herp  |
| 68 | 50.7 | 11   | 1 | TKNA_GADMO | P28498 gadus morhu  |
| 69 | 50.7 | 55   | 1 | ATP8_DICLA | Q36362 dicentrarch  |
| 70 | 50.7 | 187  | 1 | P18_LEITA  | Q25423 leishmania   |
| 71 | 50.7 | 257  | 1 | IOD1_MOUSE | Q61153 mus musculu  |
| 72 | 50.7 | 318  | 1 | MSBB_HAETN | P44567 haemophilus  |
| 73 | 50.7 | 330  | 1 | Y355_SYNY3 | P74435 synechocyst  |
| 74 | 50.7 | 359  | 1 | WN5B_HUMAN | Q9n1j7 homo sapien  |
| 75 | 50.7 | 359  | 1 | WN5B_MOUSE | P22726 mus musculu  |
| 76 | 50.7 | 360  | 1 | WN5C_XENLA | P33945 xenopus lae  |
| 77 | 50.7 | 365  | 1 | FUT3_BOVIN | Q11126 bos taurus   |
| 78 | 50.7 | 365  | 1 | WN5A_HUMAN | P41221 homo sapien  |
| 79 | 50.7 | 379  | 1 | CYB_ASTPE  | Q33818 asterina pe  |
| 80 | 50.7 | 379  | 1 | WN5A_MOUSE | P22725 mus musculu  |
| 81 | 50.7 | 379  | 1 | WN5A_MOUSE | Q9qxq7 rattus norv  |
| 82 | 50.7 | 380  | 1 | CYB_BRAFL  | O47431 branchiosto  |
| 83 | 50.7 | 380  | 1 | CYB_BRAFL  | P92472 branchiosto  |
| 84 | 50.7 | 380  | 1 | WN5A_XENLA | P31286 xenopus lae  |
| 85 | 50.7 | 381  | 1 | CYB_CARPL  | P34866 carcharhinu  |
| 86 | 50.7 | 381  | 1 | CYB_CHOCR  | P48875 chondrus cr  |
| 87 | 50.7 | 393  | 1 | CYB_PRIGL  | P34873 prionace gl  |
| 88 | 50.7 | 398  | 1 | CIW4_MOUSE | Q9nyg8 homo sapien  |
| 89 | 50.7 | 473  | 1 | CISY_SCHPO | O88454 mus musculu  |
| 90 | 50.7 | 522  | 1 | CPF4_RAT   | Q10306 schizosacch  |
| 91 | 50.7 | 1238 | 1 | DPOG_MOUSE | P51869 rattus norv  |
| 92 | 50.7 | 1507 | 1 | Y056_HUMAN | P42695 homo sapien  |

ALIGNMENTS

|          |                                               |           |      |        |
|----------|-----------------------------------------------|-----------|------|--------|
| RESULT 1 | TKNA_HORSE                                    | STANDARD; | PRT; | 11 AA. |
| ID       | TKNA_HORSE                                    |           |      |        |
| AC       | P01290;                                       |           |      |        |
| DT       | 21-JUL-1986 (Rel. 01, Created)                |           |      |        |
| DT       | 21-JUL-1986 (Rel. 01, Last sequence update)   |           |      |        |
| DT       | 30-MAY-2000 (Rel. 39, Last annotation update) |           |      |        |
| DE       | SUBSTANCE P.                                  |           |      |        |
| GN       | TAC1 OR NKNA OR TAC2 OR NKA.                  |           |      |        |

OS Equus caballus (Horse), and Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 OX NCBI\_TaxID=9796, 10141;  
 RN SEQUENCE.  
 RP SPECIES=Horse;  
 RC Studer R.O., Trzeciak A., Lergier W.;  
 RA "Isolation and amino-acid sequence of substance P from horse  
 RT intestine";  
 RL Helv. Chim. Acta 56:860-866(1973).  
 RN [2]  
 RP SEQUENCE.  
 RC SPECIES=C.porcellus;  
 RX MEDLINE=90044685; PubMed=2478925;  
 RA Murphy R.;  
 RT "Primary amino acid sequence of guinea-pig substance P";  
 RL Neuropeptides 14:105-110(1989).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 CC  
 DR PIR; A01558; SPHO.  
 DR InterPro; IPR003580; Protachykinin.  
 DR Pfam; PF02202; Tachykinin; 1.  
 DR SMART; SM00203; TK; 1.  
 DR PROSITE; PS00267; TACHYKININ; 1.  
 DR Tachykinin; Neuropeptide; Amidation; Neurotransmitter.  
 FT MOD\_RES 11  
 FT AMIDATION.  
 SQ SEQUENCE 11 AA; 1349 MW; 3E757FE3C9D6C6C7 CRC64;

Query Match 67.6%; Score 48; DB 1; Length 11;  
 Best Local Similarity 81.8%; Pred. No. 0.17;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFGLM 11  
 |||||:|:|  
 DB 1 RPKPQQFFGLM 11

RESULT 2  
 TKNL\_RABIT  
 ID TKNL\_RABIT STANDARD; PRT; 115 AA.  
 AC P41340;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROTACHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
 DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE GAMMA; C-TERMINAL  
 DE FLANKING PEPTIDE].  
 DE TAC1 OR NKNA OR TAC2 OR NKA.  
 GN Oryctolagus cuniculus (Rabbit).  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=93371392; PubMed=8363593;  
 RA Maegret H.J., Heitland A., Rose M., Forssmann W.G.;  
 RT "Nucleotide sequence of the rabbit gamma-preprotachykinin I cDNA";  
 RL Biochem. Biophys. Res. Commun. 193:128-131(1993).  
 RN [2]  
 RP SEQUENCE OF 72-92.  
 RA Kage R., McGregor G.P., Thim L., Conlon J.M.;  
 RT "Gamma-neuropeptide K: a peptide isolated from rabbit gut that is  
 RT derived from gamma-preprotachykinin";  
 RL Regul. Pept. 18:346-346(1987).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,

CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
 CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X62994; CAA44728.1; -  
 DR PIR; S18922; S18922.  
 DR InterPro; IPR003580; Protachykinin.  
 DR Pfam; PF02202; Tachykinin; 1.  
 DR ProDom; PD005598; Protachykinin; 1.  
 DR SMART; SM00203; TK; 2.  
 DR PROSITE; PS00267; TACHYKININ; 2.  
 DR Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
 KW Amidation; Alternative splicing; Signal; Neurotransmitter.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT PEPTIDE 20 56 SUBSTANCE P.  
 FT PEPTIDE 58 68 SUBSTANCE P.  
 FT PEPTIDE 72 92 NEUROPEPTIDE GAMMA.  
 FT PEPTIDE 83 92 NEUROKININ A.  
 FT PEPTIDE 96 111 C-TERMINAL FLANKING PEPTIDE.  
 FT MOD\_RES 68 68 AMIDATION (G-69 PROVIDE AMIDE GROUP).  
 FT MOD\_RES 92 92 AMIDATION (G-93 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 115 AA; 13370 MW; 5EC76F7C9B10E1C6 CRC64;

Query Match 67.6%; Score 48; DB 1; Length 115;  
 Best Local Similarity 81.8%; Pred. No. 1.4;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFGLM 11  
 |||||:|:|  
 DB 58 RPKPQQFFGLM 68

RESULT 3  
 TKNL\_HUMAN  
 ID TKNL\_HUMAN STANDARD; PRT; 129 AA.  
 AC P20366; Q00072; O60600; O60601;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROTACHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
 DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE  
 DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
 DE TAC1 OR NKNA OR TAC2 OR NKA.  
 GN Homo sapiens (Human).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM BETA).  
 RX MEDLINE=87030957; PubMed=3770210;  
 RA Harnar A.J., Armstrong A., Pascall J.C., Chapman K., Rosie F.,  
 RA Curtis A., Goring J., Edwards C.R.W., Fink G.;  
 RT "cDNA sequence of human beta-preprotachykinin, the common precursor  
 RT to substance P and neurokinin A";  
 RL FEBS Lett. 208:67-72(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM BETA).  
 RC TISSUE=Brain;  
 RA Tan A., Too H.P.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

[3]
SEQUENCE OF 36-122 FROM N.A. (ISOFORMS BETA AND GAMMA).
TISSUE-Testis;
MEDLINE=91209287; PubMed=1708336;
RA Chiwakata C., Brackmann B., Hunt N., Davidoff M., Schulze W.,
RA Iveli R.;
"tachykinin (substance-P) gene expression in Leydig cells of the
human and mouse testis.";
RL Endocrinology 128:2441-2448(1991).
[4]
SEQUENCE OF 98-107.
RA MEDLINE=87275962; PubMed=3038549;
RA Theodorsson-Norheim E., Joernvall H., Andersson M., Norheim I.,
RA Oeberg K., Jacobsson G.;
"Isolation and characterization of neurokinin A, neurokinin A(3-10)
and neurokinin A(4-10) from a neutral water extract of a metastatic
ileal carcinoid tumour.";
RL Eur. J. Biochem. 166:693-697(1987).
[5]
SEQUENCE OF 36-118 FROM N.A. (ISOFORM ALPHA).
TISSUE-Blood, and Brain;
RA Lai J.P., Douglas S.D., Rappaport E., Wu J.M., Ho W.Z.;
RA "Identification of a delta isoform of preprotachykinin mRNA in human
mononuclear phagocytes and lymphocytes.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
[6]
SEQUENCE OF 111-126.
RA TISSUE-Adrenal medulla;
RA MEDLINE=91133994; PubMed=2284201;
RA McGregor G.P., Conlon J.M.;
RA "Characterization of the C-terminal flanking peptide of human
beta-preprotachykinin.";
RL Peptides 11:907-910(1990).
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH
CC MUSCLES.
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),
CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).
-----
EMBL; X54469; CAA38351.1; -
DR EMBL; U37529; AAA79195.1; -
DR EMBL; M68906; AAA60159.1; -
DR EMBL; M68907; AAA60160.1; -
DR EMBL; AF050656; AAC15702.1; -
DR EMBL; AF050658; AAC15704.1; -
DR PIR; A24805; A24805.
DR PIR; S00069; S00069.
DR MIM; 162320; -
DR InterPro; IPR003580; Protachykinin.
DR InterPro; IPR002040; Tachykinin.
DR Pfam; PF02202; Tachykinin; 1.
DR ProDom; PD005598; Protachykinin; 1.
DR SMART; SM00203; TK; 2.
DR PROSITE; PS00267; TACHYKININ; 2.
KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;
KW Amidation; Alternative splicing; Signal; Neurotransmitter.
FT SIGNAL 1 19
FT PROPEP 20 56
FT PEPTIDE 58 68
FT PEPTIDE 72 107
FT PEPTIDE 72 73
FT PEPTIDE 89 107
FT PEPTIDE 98 107

PEPTIDE 111 126  
FT MOD\_RES 68  
FT MOD\_RES 107 107  
FT VARSPLIC 74 88  
FT VARSPLIC 97 114  
FT VARSPLIC 115 115  
FT CONFLICT 87 87  
SQ SEQUENCE 129 AA; 15003 MW; 51412C1692368DE4 CRC64;  
L -> P (IN REF. 4).  
Query Match 67.6%; Score 48; DB 1; Length 129;  
Best Local Similarity 81.8%; Pred No. 1.6;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 RPKQQQFWLM 11  
DB 58 RPKQQQFFGLM 68  
|||||: ||  
RESULT 4  
TKNL\_BOVIN  
ID TKNL\_BOVIN STANDARD; PRT; 130 AA.  
AC P01289; P01291; P04091; P20773;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DE 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROTACHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE  
DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
GN TACI OR NKNA OR TAC2 OR NKA.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM BETA).  
RX MEDLINE=85086245; PubMed=6083453;  
RA Nawa H., Kotani H., Nakanishi S.;  
RT "Tissue-specific generation of two preprotachykinin mRNAs from one  
RT gene by alternative RNA splicing.";  
RL Nature 312:729-734(1984).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).  
RX MEDLINE=84039802; PubMed=6195531;  
RA Nawa H., Hirose T., Takashima H., Inayama S., Nakanishi S.;  
RT "Nucleotide sequences of cloned cDNAs for two types of bovine brain  
RT substance P precursor.";  
RL Nature 306:32-36(1983).  
RN [3]  
RP SEQUENCE OF 36-122 FROM N.A. (ISOFORMS BETA AND GAMMA).  
RX TISSUE-Hypothalamus;  
RX MEDLINE=91209287; PubMed=1708336;  
RA Chiwakata C., Brackmann B., Hunt N., Davidoff M., Schulze W.,  
RA Iveli R.;  
RT "Tachykinin (substance-P) gene expression in Leydig cells of the  
RT human and mouse testis.";  
RL Endocrinology 128:2441-2448(1991).  
CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -!- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X00075; CAA24939.1; -;  
CC EMBL; X00075; CAA24940.1; -;  
CC EMBL; X00075; CAA24941.1; -;  
CC EMBL; X00075; CAA24942.1; -;  
CC EMBL; X00076; CAA24943.1; ALT\_SEQ.  
CC EMBL; X00076; CAA24943.1; -;  
CC EMBL; X02351; CAA26206.1; -;  
CC EMBL; X01396; CAA26206.1; JOINED.  
CC EMBL; X01397; CAA26206.1; JOINED.  
CC EMBL; X01398; CAA26206.1; JOINED.  
CC EMBL; X01399; CAA26206.1; JOINED.  
CC EMBL; X01400; CAA26206.1; JOINED.  
CC EMBL; M68911; AAA30724.1; -;  
CC EMBL; M68912; AAA30725.1; -;  
CC PIR; A01557; SPBOA.  
CC PIR; A05093; SPBOB.  
CC PIR; A05093; SPBOB.  
CC PIR; B25067; B25067.  
CC InterPro; IPR003580; Protachykinin.  
CC InterPro; IPR002040; Tachykinin.  
CC Pfam; PF02202; Tachykinin; 1.  
CC ProDom; PD005598; Protachykinin; 1.  
CC SMART; SM00203; TK; 2.  
CC PROSITE; PS00267; TACHYKININ; 2.  
CC Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
CC Amidation; Alternative splicing; Signal; Neurotransmitter.  
CC SIGNAL 1 19 POTENTIAL.  
CC PROPEP 20 56 SUBSTANCE P.  
CC PEPTIDE 58 68 NEUROPEPTIDE K.  
CC PEPTIDE 72 107 NEUROPEPTIDE GAMMA 1ST PART.  
CC PEPTIDE 72 73 NEUROPEPTIDE GAMMA 2ND PART.  
CC PEPTIDE 89 107 NEUROPEPTIDE A.  
CC PEPTIDE 111 126 C-TERMINAL FLANKING PEPTIDE (POTENTIAL).  
CC MOD\_RES 68 68 AMIDATION (G-69 PROVIDE AMIDE GROUP).  
CC MOD\_RES 107 107 AMIDATION (G-108 PROVIDE AMIDE GROUP).  
CC VARSPPLIC 74 88 MISSING (IN ISOFORM GAMMA AND ISOFORM DELTA).  
CC VARSPPLIC 97 114 MISSING (IN ISOFORM ALPHA AND ISOFORM DELTA).  
CC VARSPPLIC 115 115 V -> M (IN ISOFORM ALPHA AND ISOFORM DELTA).  
CC CONFLICT 121 121 V -> A (IN REF. 3).  
CC SEQUENCE 130 AA; 15076 MW; CE2A28572305DEB7 CRC64;

Query Match 67.6%; Score 48; DB 1; Length 130;  
Best Local Similarity 81.8%; Pred. No. 1.6;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
Db 58 RPKPQQWFGLM 68  
|||||:|

RESULT 5

TKNL\_MESAU  
ID TKNL\_MESAU STANDARD; PRT; 130 AA.  
AC Q60541; P49110;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROTACHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE  
DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
GN TAC1 OR NKNA OR TAC2 OR NKA.  
OS Mesocricetus auratus (Golden hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Mesocricetus.

OX NCBI\_TaxID=10036;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS BETA AND GAMMA).  
RC STRAIN=AURA; TISSUE=Brain;  
RA Heitland A., Kruhofer M., Juergen Maegert H.J., Forssmann W.G.;  
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
CC MUSCLES.  
CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X80662; CAA56691.1; -;  
CC EMBL; X80663; CAA56692.1; -;  
CC InterPro; IPR003580; Protachykinin.  
CC InterPro; IPR002040; Tachykinin.  
CC Pfam; PF02202; Tachykinin; 1.  
CC ProDom; PD005598; Protachykinin; 1.  
CC SMART; SM00203; TK; 2.  
CC PROSITE; PS00267; TACHYKININ; 2.  
CC Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
CC Amidation; Alternative splicing; Signal; Neurotransmitter.  
CC SIGNAL 1 19 POTENTIAL.  
CC PROPEP 20 56 SUBSTANCE P.  
CC PEPTIDE 58 68 NEUROPEPTIDE K.  
CC PEPTIDE 72 107 NEUROPEPTIDE GAMMA 1ST PART.  
CC PEPTIDE 72 73 NEUROPEPTIDE GAMMA 2ND PART.  
CC PEPTIDE 89 107 NEUROPEPTIDE A.  
CC PEPTIDE 111 126 C-TERMINAL FLANKING PEPTIDE (POTENTIAL).  
CC MOD\_RES 68 68 AMIDATION (G-69 PROVIDE AMIDE GROUP).  
CC MOD\_RES 107 107 AMIDATION (G-108 PROVIDE AMIDE GROUP).  
CC VARSPPLIC 74 88 MISSING (IN ISOFORM GAMMA);  
CC SEQUENCE 130 AA; 14907 MW; CC92E9371A646F2E CRC64;

Query Match 67.6%; Score 48; DB 1; Length 130;  
Best Local Similarity 81.8%; Pred. No. 1.6;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 RPKPQQWFWM 11  
Db 58 RPKPQQWFGLM 68  
|||||:|

RESULT 6

TKNL\_MOUSE  
ID TKNL\_MOUSE STANDARD; PRT; 130 AA.  
AC P41539; Q00073;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE PROTACHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE  
DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
GN TAC1 OR NKNA OR TAC2 OR NKA.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OC NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM BETA).  
RC STRAIN=ICR; TISSUE=Brain;

RA KAKO K., MUNEKATA E., HOSAKA M., MURAKAMI K., NAKAYAMA K.;  
 RT "Cloning and sequence analysis of mouse cDNAs encoding  
 RT preprotachykinin A and B";  
 RL Biomed. Res. 14:253-259(1993).  
 RN [2]  
 RP SEQUENCE OF 36-122 FROM N.A. (ISOFORMS BETA AND GAMMA).  
 RC TISSUE=Brain;  
 RX MEDLINE=91209287; PubMed=1708336;  
 RA Chivakata C., Brackmann B., Hunt N., Davidoff M., Schulze W.,  
 RA Ivell R.;  
 RT "Tachykinin (substance-P) gene expression in Leydig cells of the  
 RT human and mouse testis";  
 RL Endocrinology 128:2441-2448(1991).  
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
 CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DE EMBL; D17584; BAA04508.1; -;  
 DR EMBL; M68908; AAA39969.1; -;  
 DR EMBL; M68909; AAA39970.1; -;  
 DR MGD; MGI:98474; Tactl.  
 DR InterPro; IPR003580; Protachykinin.  
 DR InterPro; IPR002040; Tachykinin.  
 DR Pfam; PF02202; Tachykinin; 1.  
 DR ProDom; PD005598; Protachykinin; 1.  
 DR SMART; SM00203; TK; 2.  
 DR PROSITE; PS00267; TACHYKININ; 2.  
 KW Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
 KW Amidation; Alternative splicing; Signal; Neurotransmitter.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT PROPEP 20 56 SUBSTANCE P.  
 FT PEPTIDE 58 68 NEUROPEPTIDE K.  
 FT PEPTIDE 72 107 NEUROPEPTIDE GAMMA 1ST PART.  
 FT PEPTIDE 72 73 NEUROPEPTIDE GAMMA 2ND PART.  
 FT PEPTIDE 89 107 NEUROKININ A.  
 FT PEPTIDE 98 107 NEUROKININ A.  
 FT PEPTIDE 111 126 C-TERMINAL FLANKING PEPTIDE (POTENTIAL).  
 FT MOD\_RES 68 68 AMIDATION (G-69 PROVIDE AMIDE GROUP).  
 FT MOD\_RES 107 107 AMIDATION (G-108 PROVIDE AMIDE GROUP).  
 FT VARSPIC 74 88 MISSING (IN ISOFORM GAMMA).  
 SQ SEQUENCE 130 AA; 15045 MW; 7BE8DA15FDE72FF8 CRC64;  
 Query Match 67.6%; Score 48; DB 1; Length 130;  
 Best Local Similarity 81.8%; Pred. No. 1.6;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RPKPQQFWLM 11  
 Db 58 RPKPQQFFGLM 68  
 RESULT 7  
 TKNI\_RAT STANDARD; PRT; 130 AA.  
 AC P06767; P08856; P22356;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PROTACHYKININ 1 PRECURSOR (PPT) [CONTAINS: SUBSTANCE P; NEUROKININ A  
 DE (NKA) (SUBSTANCE K) (NEUROMEDIN L); NEUROPEPTIDE K (NPK); NEUROPEPTIDE

DE GAMMA; C-TERMINAL FLANKING PEPTIDE].  
 GN TAC1 OR NKNA OR TAC2 OR NKA.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA AND GAMMA).  
 RX MEDLINE=90331040; PubMed=1695945;  
 RA Carter M.S., Krause J.E.;  
 RT "Structure, expression, and some regulatory mechanisms of the rat  
 RT preprotachykinin gene encoding substance P, neurokinin A,  
 RT neuropeptide K, and neuropeptide gamma";  
 RL J. Neurosci. 10:2203-2214(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS ALPHA; BETA AND GAMMA).  
 RX MEDLINE=87118268; PubMed=2433692;  
 RA Krause J.E., Chirgwin J.M., Carter M.S., Xu Z.S., Hershey A.D.;  
 RT "Three rat preprotachykinin mRNAs encode the neuropeptides substance  
 RT P and neurokinin A";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:881-885(1987).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM GAMMA).  
 RX MEDLINE=87025808; PubMed=2429656;  
 RA Kawaguchi Y., Hoshimaru M., Nawa H., Nakanishi S.;  
 RT "Sequence analysis of cloned cDNA for rat substance P precursor:  
 RT existence of a third substance P precursor";  
 RL Biochem. Biophys. Res. Commun. 139:1040-1046(1986).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM DELTA).  
 RC TISSUE=Dorsal root ganglion;  
 RX MEDLINE=91085565; PubMed=1702066;  
 RA Harmar A.J., Hyde V., Chapman K.E.;  
 RT "Identification and cDNA sequence of delta-preprotachykinin, a fourth  
 RT splicing variant of the rat substance P precursor";  
 RL FEBS Lett. 275:22-24(1990).  
 RN [5]  
 RP SEQUENCE OF 1-41 FROM N.A.  
 RX MEDLINE=93192337; PubMed=8448217;  
 RA Chapman K.E., Lyons V., Harmar A.J.;  
 RT "The sequence of 5' flanking DNA from the rat preprotachykinin gene;  
 RT analysis of putative transcription factor binding sites";  
 RL Biochim. Biophys. Acta 1172:361-363(1993).  
 CC -1- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOKE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -1- ALTERNATIVE PRODUCTS: 4 ISOFORMS; ALPHA, BETA (SHOWN HERE),  
 CC GAMMA AND DELTA; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DE EMBL; M34162; AAA41926.1; -;  
 DR EMBL; M34159; AAA41926.1; JOINED.  
 DR EMBL; M34160; AAA41926.1; JOINED.  
 DR EMBL; M34161; AAA41926.1; JOINED.  
 DR EMBL; M34164; AAA41925.1; -;  
 DR EMBL; M34183; AAA41929.1; -;  
 DR EMBL; M15191; AAA41928.1; -;  
 DR EMBL; M14312; AAA41927.1; -;  
 DR EMBL; L07328; AAA41924.1; -;  
 DR EMBL; X56306; CAA39752.1; -;  
 DR PIR; A26590; A26590.  
 DR PIR; B26590; B26590.  
 DR PIR; C26590; C26590.  
 DR PIR; A37163; A37163.

DR PIR; S12958; S12958.  
 DR InterPro; IPR003580; Protachykinin.  
 DR InterPro; IPR002040; Tachykinin.  
 DR Pfam; PF02202; Tachykinin; 1.  
 DR ProDom; PD005598; Protachykinin; 1.  
 DR SMART; SM00203; TK; 2.  
 DR PROSITE; PS00267; TACHYKININ; 2.  
 DR Tachykinin; Neuropeptide; Cleavage on pair of basic residues;  
 KW Amidation; Alternative splicing; Signal; Neurotransmitter.  
 FT SIGNAL 1 19 POTENTIAL.  
 FT PROPEP 20 56 POTENTIAL.  
 FT PEPTIDE 58 68 SUBSTANCE P.  
 FT PEPTIDE 72 107 NEUROPEPTIDE K.  
 FT PEPTIDE 72 73 NEUROPEPTIDE GAMMA 1ST PART.  
 FT PEPTIDE 89 107 NEUROPEPTIDE GAMMA 2ND PART.  
 FT PEPTIDE 98 107 NEUROKININ A.  
 FT PEPTIDE 111 126 C-TERMINAL FLANKING PEPTIDE (POTENTIAL).  
 FT MOD\_RES 68 68 AMIDATION (G-69 PROVIDE AMIDE GROUP).  
 FT MOD\_RES 107 107 AMIDATION (G-108 PROVIDE AMIDE GROUP).  
 FT VARSPPLIC 74 88 MISSING (IN ISOFORM GAMMA AND ISOFORM DELTA).  
 FT VARSPPLIC 97 114 MISSING (IN ISOFORM ALPHA AND ISOFORM DELTA).  
 FT VARSPPLIC 115 115 V -> M (IN ISOFORM ALPHA AND ISOFORM DELTA).  
 FT SEQUENCE 130 AA; 15001 MW; B22EFE860DCCD75A CRC64;

Query Match 67.6%; Score 48; DB 1; Length 130;  
 Best Local Similarity 81.8%; Pred. No. 1.6;  
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPKPQQWFNLM 11  
 |||||:|:|  
 Db 58 RRPQQQFFGLM 68

RESULT 8  
 TKNA\_CHICK STANDARD; PRT; 11 AA.  
 ID TKNA\_CHICK  
 AC F19850;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE SUBSTANCE P.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE  
 RC TISSUE=Intestine; PubMed=2452461;  
 RX MEDLINE=88204263; PubMed=2452461;  
 RA Conlon J.M., Katsoulis S., Schmidt W.E., Thim L.;  
 RT "[Arg3]substance P and neurokinin A from chicken small intestine.";  
 RL Regul. Pept. 20:171-180(1988).  
 CC -!- FUNCTION: TACHYKININS ARE ACTIVE PEPTIDES WHICH EXCITE NEURONS,  
 CC EVOLVE BEHAVIORAL RESPONSES, ARE POTENT VASODILATORS AND  
 CC SECRETAGOGUES, AND CONTRACT (DIRECTLY OR INDIRECTLY) MANY SMOOTH  
 CC MUSCLES.  
 CC -!- SIMILARITY: BELONGS TO THE TACHYKININ FAMILY.  
 DR PIR; JN0023; JN0023.  
 DR InterPro; IPR003580; Protachykinin.  
 DR InterPro; IPR002040; Tachykinin.  
 DR Pfam; PF02202; Tachykinin; 1.  
 DR SMART; SM00203; TK; 1.  
 DR PROSITE; PS00267; TACHYKININ; 1.  
 DR Tachykinin; Neuropeptide; Amidation; Neurotransmitter.  
 KW MOD\_RES 11 11 AMIDATION  
 FT SEQUENCE 11 AA; 1377 MW; 21487FE3C9D6C6C7 CRC64;

Query Match 63.4%; Score 45; DB 1; Length 11;  
 Best Local Similarity 72.7%; Pred. No. 0.44;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 RPKPQQWFNLM 11  
 ||:||||:|:|  
 Db 1 RRPQQQFFGLM 11  
 RESULT 9  
 CYB\_RHOCA STANDARD; PRT; 436 AA.  
 ID CYB\_RHOCA  
 AC P08502; P07057;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE CYTOCHROME B.  
 GN PETB OR CYTB.  
 OS Rhodobacter capsulatus (Rhodopseudomonas capsulata).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;  
 OC Rhodobacter.  
 OC NCBI\_TaxID=1061;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SB1003;  
 RX MEDLINE=88011223; PubMed=2821268;  
 RA Davidson E., Daldal F.;  
 RT "Primary structure of the bcl complex of Rhodopseudomonas capsulata.  
 RT Nucleotide sequence of the pet operon encoding the Rieske cytochrome  
 RT b, and cytochrome c1 apoproteins.";  
 RL J. Mol. Biol. 195:13-24(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GA;  
 RX MEDLINE=86136096; PubMed=3004982;  
 RA Gabellini N., Seibald W.;  
 RT "Nucleotide sequence and transcription of the fbc operon from  
 RT Rhodopseudomonas sphaeroides. Evaluation of the deduced amino acid  
 RT sequences of the Fes protein, cytochrome b and cytochrome c1.";  
 RL Eur. J. Biochem. 154:569-579(1986).  
 RN [3]  
 RP CORRECTION OF ORGANISM GIVEN IN REF. 2.  
 RX MEDLINE=88011233; PubMed=2821272;  
 RA Davidson E., Daldal F.;  
 RT "fbc operon, encoding the Rieske Fe-S protein cytochrome b, and  
 RT cytochrome c1 apoproteins previously described from Rhodopseudomonas  
 RT sphaeroides, is from Rhodopseudomonas capsulata.";  
 RL J. Mol. Biol. 195:25-29(1987).  
 RN [4]  
 RP MUTATIONS CONFERRING RESISTANCE TO QUINOL OXIDATION INHIBITORS.  
 RX MEDLINE=90076115; PubMed=2556259;  
 RA Daldal F., Tokito M.K., Davidson E., Faham M.;  
 RT "Mutations conferring resistance to quinol oxidation (Qz) inhibitors  
 RT of the cyt bcl complex of Rhodobacter capsulatus.";  
 RL EMBO J. 8:3951-3961(1989).  
 RN [5]  
 RP MUTAGENESIS.  
 RC STRAIN=WT1131;  
 RX MEDLINE=91105061; PubMed=2176897;  
 RA Robertson D.E., Daldal F., Dutton P.L.;  
 RT "Mutants of ubiquinol-cytochrome c2 oxidoreductase resistant to Qo  
 RT site inhibitors: consequences for ubiquinone and ubiquinol affinity  
 RT and catalysis";  
 RL Biochemistry 29:11249-11260(1990).  
 CC -!- FUNCTION: COMPONENT OF THE UBIQUINOL-CYTOCHROME C REDUCTASE  
 CC COMPLEX (COMPLEX III OR CYTOCHROME B-C1 COMPLEX), WHICH IS A  
 CC RESPIRATORY CHAIN THAT GENERATES AN ELECTROCHEMICAL POTENTIAL  
 CC COUPLED TO ATP SYNTHESIS.  
 CC -!- COFACTOR: TWO HEME GROUPS (B562 AND B566) WHICH ARE NOT COVALENTLY  
 CC BOUND TO THE PROTEIN.  
 CC -!- SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,  
 CC CYTOCHROME C1 AND THE RIESKE PROTEIN.  
 CC -!- SIMILARITY: BELONGS TO THE CYTOCHROME B FAMILY.

```
RESULT 1
AAP61480
ID AAP61480 standard; peptide; 11 AA.
XX
AC AAP61480;
XX
DT 22-AUG-1991 (first entry)
XX
DE Sequence of undeca peptide substance P1.
XX
KW Hypertension therapy; sleep disorder; anti-stress agent.
XX
FH Key
FT Misc-difference 11
FT /label= Met-NH2
XX
PN DD229593-A.
XX
PD 13-NOV-1985.
XX
PF 28-NOV-1984; 84DD-0269954.
XX
PR 28-NOV-1984; 84DD-0269954.
XX
PA (DEAK ) AKAD WISSENSCHAFT DDR.
XX
PI Oehme P, Hecht K, Wachtel E, Roske I, Kolometsewa IA;
PI Airpetjanz M, Bliener M, Vogt WE, Hilse H, Gores E, Poppei M;
PI Nieber K, Bergmann J;
XX
DR WPI; 1986-069587/11.
XX
PT Cpd. having N-terminal sequences of undeca:peptide substance P -
PT are medicinal agents with anti-stress activity
XX
PS Claim 1; Page 1; 15pp; German.
XX
CC The inventors claim an antistress compound which contains the N-
CC terminal SQ of AAP61480, pref. Arg-Pro-Lys-Pro-X (X= COOH or NH2).
CC Compared with the full undecapeptide they have much reduced
CC side effects (acute hypotension, spastic effects on the ileum and
CC histamine release from peritoneal mast cells).
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 7; Length 11;
Best Local Similarity 100.0%; Pred. NO. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQQFFGLM 11
Db 1 rpkpgqffgllm 11

RESULT 2
AAP80312
ID AAP80312 standard; protein; 11 AA.
XX
AC AAP80312;
XX
DT 14-SEP-1990 (first entry)
XX
DE Sequence of neuropeptide substance P which binds with polypeptide
DE receptor for bombesin type polypeptides.
XX
KW Spantide; neuropeptide; polypeptide receptor; cancer diagnosis;
KW cancer therapy; Swiss 3T3 cells; bombesin type polypeptides;
XX
OS Swiss 3T3 cells.
XX
FH Key
FH Location/Qualifiers

Query Match 100.0%; Score 61; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. NO. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQQFFGLM 11
Db 1 rpkpgqffgllm 11

RESULT 3
AAR13162
ID AAR13162 standard; Protein; 11 AA.
XX
AC AAR13162;
XX
DT 10-OCT-1991 (first entry)
XX
DE Sialic acid-bonded polypeptide (2).
XX
KW Sialic acid; cataract; immune disorder.
XX
OS Synthetic.
XX
FH Key
FH Location/Qualifiers
FH Modified-site 1
FH /note= "N-terminally glycosylated by 5-acetamido-
FT 2,4,7,8,9-penta-O-acetyl-3,5-deoxy-beta-
FT D-glycero-D-galactononulopyranosyl"
```

PN JP03151398-A.  
XX 27-JUN-1991.  
XX  
PF 06-NOV-1989; 89JP-0288560.  
XX  
PR 06-NOV-1989; 89JP-0288560.  
XX  
PA (MECT-) MECT KK.  
XX  
DR WPI; 1991-233839/32.  
XX  
XX New sialic acid derivs. bonded to physiologically active  
PT polypeptide - for treatment of cataracts, immune disorders etc.  
PT with prolonged half-life  
XX  
PS Example 4; Page 6; 7pp; Japanese.  
XX  
CC The prod. has prolonged half-life and is used as a pharmaceutical  
CC for treatment of various diseases, such as cataract and immune  
CC disorders. It comprises a peptide, N-terminally glycosylated by  
CC (opt. acetylated) sialic acid.  
CC See also AAR12932, AAR13162 and AAR13201.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 100.0%; Score 61; DB 12; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11  
| | | | | | | | | |  
Db 1 rpkgqgffglm 11

RESULT 4  
AAR11854  
ID AAR11854 standard; peptide; 11 AA.  
XX  
AC AAR11854;  
XX  
DT 09-JUL-1991 (first entry)  
XX  
DE Undecapeptide substance P.  
XX  
KW Undecapeptide; pharmaceutical; stress; sleep.  
XX  
OS Synthetic.  
XX  
PN DD285097-A.  
XX  
PD 05-DEC-1990.  
XX  
PF 21-JUN-1989; 89DD-0329831.  
XX  
PR 21-JUN-1989; 89DD-0329831.  
XX  
PA (DEAK ) INST WIRKSTOFF AKAD.  
XX  
FA (FAF ) VEB CHEM BITTERFELD.  
XX  
PI Beyermann M, Bienert M, Egler H, Haupke K, Krause E;  
XX Schwarz J, Walz H;  
XX WPI; 1991-133498/19.  
DR  
XX Undeca-peptide substance pharmaceutical intermediate prepn. - by  
PT forming di-peptide between nitro-arginine and proline and  
PT reacting with polymer-bound non-peptide  
XX  
PS Calim 1; Page 1; 8pp; German.  
XX  
CC The peptide is prepared by solid phase synthesis.

CC It can be used in the preparation of pharmaceuticals which can be  
CC used to treat certain stress-induced disturbances of the sleep  
CC profile.  
XX  
SQ Sequence 11 AA;  
XX

Query Match 100.0%; Score 61; DB 12; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11  
| | | | | | | | | |  
Db 1 rpkgqgffglm 11

RESULT 5  
AAR21938  
ID AAR21938 standard; Protein; 11 AA.  
XX  
AC AAR21938;  
XX  
DT 25-JUN-1992 (first entry)  
XX  
DE Substance P [Me-Leu 10].  
XX  
KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 10  
FT /label= OTHER  
FT /note= "OTHER - Me-Leu"  
XX  
XX WO9202248-A.  
XX  
XX 20-FEB-1992.  
XX  
XX 29-JUL-1991; 91WO-US05323.  
XX  
XX 27-JUL-1990; 90US-0559173.  
XX  
XX (CHIL-) CHILDRENS MED CENT.  
XX  
XX Yankner BA;  
XX  
XX WPI; 1992-079804/10.  
XX  
XX Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX  
XX Claim 10; Page 21; 35pp; English.  
XX  
XX The peptide is the tachykinin agonist substance P with Me-Leu  
CC substituted at position 10. The peptide was synthesised  
CC by standard solid phase synthesis. Neuronal accumulation of  
CC beta-amyloid may be treated by administration of tachykinin  
CC agonists. The peptides can reduce the neurotoxic effects of a  
CC beta-amyloid related polypeptide on cultured neurons. The peptide  
CC and its analogues are useful for controlling diseases characterised  
CC by beta amyloid accumulation in the brain such as Alzheimer's  
CC disease and Down's syndrome.  
CC See also AAR21932-75.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 100.0%; Score 61; DB 13; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 RPKPQOFFGLM 11  
 |||||  
 Db 1 rpkpqffglm 11

RESULT 6  
 AAR21942  
 ID AAR21942 standard; Protein; 11 AA.  
 XX  
 AC AAR21942;  
 XX  
 DT 25-JUN-1992 (first entry)  
 XX  
 DE Substance P [MeMet 11].  
 XX  
 KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
 KW syndrome; hereditary cerebral haemorrhage.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 11  
 FT /label= OTHER  
 FT /note= "OTHER = Methyl Methionine"  
 XX  
 PN W09202248-A.  
 XX  
 PD 20-FEB-1992.  
 XX  
 PF 29-JUL-1991; 91WO-US05323.  
 XX  
 PR 27-JUL-1990; 90US-0559173.  
 XX  
 PA (CHIL-) CHILDRENS MED CENT.  
 XX  
 PI Yankner BA;  
 XX  
 DR WPI; 1992-079804/10.  
 XX  
 XX Treatment of neuronal accumulation of beta-amyloid - using  
 PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
 PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
 XX  
 PS Claim 10; Page 21; 35pp; English.  
 XX  
 CC The peptide is the tachykinin agonist substance P with a methyl  
 CC methionine residue substituted at position 11. The peptide was  
 CC synthesised by standard solid phase synthesis. Neuronal  
 CC accumulation of beta-amyloid may be treated by administration of  
 CC tachykinin agonists. The peptide can reduce the neurotoxic effects  
 CC of a beta-amyloid related polypeptide on cultured neurons. The  
 CC peptide and its analogues are useful for controlling diseases  
 CC characterised by beta amyloid accumulation in the brain such as  
 CC Alzheimer's disease and Down's syndrome.  
 XX  
 SQ See also AAR21932-75.  
 XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 13; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOFFGLM 11  
 |||||  
 Db 1 rpkpqffglm 11

RESULT 7  
 AAR21946  
 ID AAR21946 standard; Protein; 11 AA.  
 XX

AC AAR21946;  
 XX  
 DT 25-JUN-1992 (first entry)  
 XX  
 DE Substance P [Me-Phe 8].  
 XX  
 KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
 KW syndrome; hereditary cerebral haemorrhage.  
 XX  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 8  
 FT /label= OTHER  
 FT /note= "OTHER = Methyl phenylalanine"  
 XX  
 PN W09202248-A.  
 XX  
 PD 20-FEB-1992.  
 XX  
 PF 29-JUL-1991; 91WO-US05323.  
 XX  
 PR 27-JUL-1990; 90US-0559173.  
 XX  
 PA (CHIL-) CHILDRENS MED CENT.  
 XX  
 PI Yankner BA;  
 XX  
 DR WPI; 1992-079804/10.  
 XX  
 XX Treatment of neuronal accumulation of beta-amyloid - using  
 PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
 PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
 XX  
 PS Claim 10; Page 21; 35pp; English.  
 XX  
 CC The peptide is the tachykinin agonist substance P with a methyl  
 CC phenylalanine residue substituted at position 8. The peptide was  
 CC synthesised by standard solid phase synthesis. Neuronal  
 CC accumulation of beta-amyloid may be treated by administration of  
 CC tachykinin agonists. The peptide can reduce the neurotoxic effects  
 CC of a beta-amyloid related polypeptide on cultured neurons. The  
 CC peptide and its analogues are useful for controlling diseases  
 CC characterised by beta amyloid accumulation in the brain such as  
 CC Alzheimer's disease and Down's syndrome.  
 XX  
 SQ See also AAR21932-75.  
 XX  
 SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 13; Length 11;  
 Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOFFGLM 11  
 |||||  
 Db 1 rpkpqffglm 11

RESULT 8  
 AAR21954  
 ID AAR21954 standard; Protein; 11 AA.  
 XX  
 AC AAR21954;  
 XX  
 DT 25-JUN-1992 (first entry)  
 XX  
 DE Substance P [Me-Gly 9].  
 XX  
 KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
 KW syndrome; hereditary cerebral haemorrhage.  
 XX  
 OS Synthetic.

```
XX Key Location/Qualifiers
FH Misc-difference 9
FT /label= OTHER
FT /note= "OTHER = Methyl glycine"
XX
XX PN W09202248-A.
XX PD 20-FEB-1992.
XX PF 29-JUL-1991; 91WO-US05323.
XX PR 27-JUL-1990; 90US-0559173.
XX PA (CHIL-) CHILDRENS MED CENT.
XX PI Yankner BA;
XX DR WPI; 1992-079804/10.
XX
XX PT Treatment of neuronal accumulation of beta-amyloid - using
XX PT tachykinin agonists e.g. substance P, physalaemin and neurokinin
XX PT B, for treating Alzheimer's disease, Downs syndrome, etc.
XX PS Claim 10; Page 22; 35pp; English.
XX
XX CC The peptide is the tachykinin agonist substance P with a methyl
XX CC glycine residue substituted at position 9. The peptide was
XX CC synthesised by standard solid phase synthesis. Neuronal
XX CC accumulation of beta-amyloid may be treated by administration of
XX CC tachykinin agonists. The peptide can reduce the neurotoxic effects
XX CC of a beta-amyloid related polypeptide on cultured neurons. The
XX CC peptide and its analogues are useful for controlling diseases
XX CC characterised by beta amyloid accumulation in the brain such as
XX CC Alzheimer's disease and Down's syndrome.
XX CC See also AAR21932-75.
XX SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 13; Length 11;
Best Local Similarity 100.0%; Pred. NO. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11
Db 1 rpqpqffglm 11

RESULT 9
AAR21962
ID AAR21962 standard; Peptide; 11 AA.
XX
XX AC AAR21962;
XX
XX DT 25-JUN-1992 (first entry)
XX
XX DE Substance P [Met Gly 6, Met (O2) 11].
XX
XX KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;
XX KW syndrome; hereditary cerebral haemorrhage.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
FT Misc-difference 6
FT /label= OTHER
FT /note= "OTHER = Methyl glycine"
XX
XX FT Misc-difference 11
XX FT /label= OTHER
XX FT /note= "OTHER = Met (O2)"
XX
XX PN W09202248-A.
```

```
XX
XX PD 20-FEB-1992.
XX
XX PF 29-JUL-1991; 91WO-US05323.
XX
XX PR 27-JUL-1990; 90US-0559173.
XX
XX PA (CHIL-) CHILDRENS MED CENT.
XX
XX PI Yankner BA;
XX
XX DR WPI; 1992-079804/10.
XX
XX PT Treatment of neuronal accumulation of beta-amyloid - using
XX PT tachykinin agonists e.g. substance P, physalaemin and neurokinin
XX PT B, for treating Alzheimer's disease, Downs syndrome, etc.
XX
XX PS Claim 10; Page 22; 35pp; English.
XX
XX CC The peptide is the tachykinin agonist, substance P with methyl
XX CC glycine substituted at position 9 and Met (O2) at position 11.
XX CC The peptide was synthesised by standard solid phase synthesis.
XX CC Neuronal accumulation of beta-amyloid may be treated by administ-
XX CC ration of tachykinin agonists. The peptide can reduce the neuro-
XX CC toxic effects of a beta-amyloid related polypeptide on cultured
XX CC neurons. The peptide and its analogues are useful for controlling
XX CC diseases characterised by beta amyloid accumulation in the brain
XX CC such as Alzheimer's disease and Down's syndrome.
XX CC See also AAR21932-75.
XX
XX SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 13; Length 11;
Best Local Similarity 100.0%; Pred. NO. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11
Db 1 rpqpqffglm 11

RESULT 10
AAR21963
ID AAR21963 standard; Peptide; 11 AA.
XX
XX AC AAR21963;
XX
XX DT 25-JUN-1992 (first entry)
XX
XX DE Substance P [p-Chloro-Phe 7,8].
XX
XX KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;
XX KW syndrome; hereditary cerebral haemorrhage.
XX
XX OS Synthetic.
XX
XX FH Key Location/Qualifiers
FT Modified-site 7
FT /label= OTHER
FT /note= "OTHER = p-Chloro-phenylalanine"
XX
XX FT Modified-site 8
XX FT /label= OTHER
XX FT /note= "OTHER = p-Chloro-phenylalanine"
XX
XX PN W09202248-A.
XX
XX PD 20-FEB-1992.
XX
XX PF 29-JUL-1991; 91WO-US05323.
XX
XX PR 27-JUL-1990; 90US-0559173.
XX
```

PA (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using

PT tachykinin agonists e.g. substance P, physalaenin and neurokinin

PT B, for treating Alzheimer's disease, Downs syndrome, etc.

XX

PS Claim 10; Page 22; 35pp; English.

XX

CC The peptide is the tachykinin agonist, substance P fragment  
 CC with p-Chloro-phenylalanine residues substituted at positions 7 and  
 CC 8. The peptide was synthesised by standard solid phase synthesis.  
 CC Neuronal accumulation of beta-amyloid may be treated by administ-  
 CC ration of tachykinin agonists. The peptide can reduce the neuro-  
 CC toxic effects of a beta-amyloid related polypeptide on cultured  
 CC neurons. The peptide and its analogues are useful for controlling  
 CC diseases characterised by beta amyloid accumulation in the brain  
 CC such as Alzheimer's disease and Down's syndrome.  
 CC See also AAR21932-75.

XX Sequence 11 AA;

Query Match 100.0%; Score 61; DB 13; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00029;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOQFFGLM 11

DB 1 rpkpqqffglm 11

RESULT 11

AAR28442

ID AAR28442 standard; peptide; 11 AA.

XX AAR28442;

DT 22-MAR-1993 (first entry)

XX Substance P.

XX NK1 receptor; tumour; malignant glioma; pheochromocytoma;

KW paraganglia; small cell lung cancer; nerve regeneration; lymphoma;

KW granuloma; Crohn's disease.

XX Synthetic.

FH Key Location/Qualifiers

FT Modified-site 11

FT /note= "amidated"

XX W09218536-A.

XX 29-OCT-1992.

XX 22-APR-1992; 92WO-US03307.

XX 22-APR-1991; 91EP-0200955.

XX (MLCW ) MALLINCKRODT MEDICAL INC.

XX Bakker WH, Hagen PM, Krenning EP, Lamberts SWJ, Visser TJ;

XX WPI; 1992-382047/46.

XX Detection and localisation of tissues with neurokinine-1 receptors

PT - for detecting and treating tumours having neurokinine-1

PT receptors e.g. malignant glioma, small cell lung cancer etc.

XX

PS Disclosure; Page 4; 22pp; English.

XX Substance P or its Tyr0 deriv. is a preferred peptide having a

CC selective affinity to neurokinine-1 receptors which (when

CC labelled with a radioactive isotope) can be used in imaging methods.

CC A generic formula for preferred peptides is AAR28441. Such peptides

CC are thus useful in diagnosis and treatment of conditions that are

CC related to NK1 receptors and in visualising NK1 receptors on certain

CC tissues. See also AAR28443-R28446.

XX

SQ Sequence 11 AA;

Query Match

Best Local Similarity 100.0%; Score 61; DB 13; Length 11;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOQFFGLM 11

DB 1 rpkpqqffglm 11

RESULT 12

AAR42646

ID AAR42646 standard; peptide; 11 AA.

XX AAR42646;

AC 19-APR-1994 (first entry)

DT 19-APR-1994 (first entry)

XX Neurokinin 1 receptor affinity-contg. peptide (Substance P).

DE Neurokinin 1; somatostatin; receptor; cytokine; growth factor;

KW hormone; intra-operativ; tumour; low energy gamma photon;

KW radionuclide.

XX Synthetic.

FH Key Location/Qualifiers

FT Modified-site 11

FT /note= "the C-terminal is amidated"

XX W09318797-A.

XX 30-SEP-1993.

XX 24-MAR-1993; 93WO-US02772.

XX 25-MAR-1992; 92EP-0200848.

XX (MLCW ) MALLINCKRODT MEDICAL INC.

XX Doedens BJ, Ensing GJ, Panek KJ;

XX WPI; 1993-320461/40.

XX Intra-operatively detecting and locating tumour tissues - using

PT specific peptide(s) labelled with low energy gamma photon

PT emitting radionuclide

XX Disclosure; Page 4; 31pp; English.

XX The method of intraoperatively detecting and locating tumoral

CC tissues makes use of peptides having selective neurokinin 1

CC receptor affinity (AAR42644; generic formula; AAR42646-R42650;

CC specific examples), peptides having selective somatostatin

CC receptor affinity (AAR42645; generic formula; AAR42651-R42660;

CC specific examples), and peptides selected from cytokines;

CC growth factors and hormones.

XX Sequence 11 AA;

Query Match 100.0%; Score 61; DB 14; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOFFGLM 11  
|||||  
Db 1 rpkpqffglm 11

RESULT 13  
AAR85243  
ID AAR85243 standard; peptide; 11 AA.  
XX  
AC AAR85243;  
XX  
DT 18-AUG-1997 (first entry)  
XX  
DE Substance P peptide.  
XX  
KW Ligand; antibody; receptor; SELEX; random library; amplification; PCR;  
KW Systematic Evolution of Ligands by Exponential enrichment; primer;  
KW polymerase chain reaction; structure; mimicry; substance P; tachykinin;  
KW neuropeptide; rheumatoid arthritis; atherosclerosis; cancer;  
KW diabetic retinopathy.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 11 /note= "contains C-terminal NH2 group"  
XX  
PN W09530775-A1.  
XX  
PD 16-NOV-1995.  
XX  
PF 03-MAY-1995; 95WO-US05600.  
XX  
PR 21-DEC-1994; 94US-0361795.  
PR 06-MAY-1994; 94US-0238863.  
PR 24-MAY-1994; 94US-0248632.  
PR 09-SEP-1994; 94US-0303362.  
PR 11-JUN-1990; 90US-0536428.  
PR 10-JUN-1991; 91US-0714131.  
PR 21-OCT-1992; 92US-0964624.  
XX  
PA (UYRE-) UNIV RES CORP.  
XX  
PI Allen P, Doudna JA, Feigon J, Gold L, Nieuwlandt D;  
PI Schneider DJ, Sullenger BA, Wecker M;  
XX  
DR WPI; 1995-404132/51.  
XX  
PT Systematic evolution of ligands by exponential enrichment - for  
PT identifying nucleic acid ligands used in the treatment of, e.g. type  
PT B insulin resistance and HIV  
XX  
PS Example 10; Fig 8; 209pp; English.  
XX  
CC The invention relates to a novel method of isolating ligands that bind  
CC to target proteins e.g. antibodies or receptors, which bind other  
CC proteins or ligands. The method, designated Systematic Evolution of  
CC Ligands by Exponential enrichment (SELEX), comprises generating a library  
CC of random oligonucleotide sequences, about 40-60 nucleotides in length,  
CC and binding these sequences to the target proteins. After removal of  
CC unbound material, the remaining bound nucleotides sequences are amplified  
CC e.g. by PCR, and the newly amplified material is bound again with the  
CC target protein. This cycle continues until a sufficiently pure  
CC oligonucleotide sequence is isolated. The method allows the isolation of  
CC oligonucleotide sequences which structurally mimic the target protein's  
CC ligand. Ligands AAR06098-130 are examples of nucleic acid ligands which  
CC bind the tachykinin-family neuropeptide Substance P (this sequence). The  
CC new ligands were split into 2 groups based on their affinities for  
CC Substance P. Class 1 ligands had binding affinities up to 2 micromolar

CC whereas class 2 ligands bound at above 2 micromolar. This sequence  
CC represents the consensus of the class 1 ligands. The ligands can be  
CC used to block the activity of Substance P and is useful in the treatment  
CC of e.g. rheumatoid arthritis, atherosclerosis, diabetic retinopathy or  
CC cancer.  
XX  
SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 16; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOFFGLM 11  
|||||  
Db 1 rpkpqffglm 11

RESULT 14  
AAR77310  
ID AAR77310 standard; peptide; 11 AA.  
XX  
AC AAR77310;  
XX  
DT 08-MAR-1996 (first entry)  
XX  
DE Substance P.  
XX  
KW Substance P; neurokinin; neurokinin receptor antagonist;  
KW sensory perception; tachykinin receptor; therapy;  
KW neurodegenerative disorder; Alzheimer's disease; demyelinating disease;  
KW multiple sclerosis; respiratory disease; ocular disease;  
KW addiction disorder; adverse immune reaction; gastrointestinal disorder;  
KW bladder function disorder; fibrosing disease; collagen disease;  
KW diagnosis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 11 /note= "amidated"  
XX  
PN US5434158-A.  
XX  
PD 18-JUL-1995.  
XX  
PF 26-APR-1994; 94US-0233487.  
XX  
PR 26-APR-1994; 94US-0233487.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Shah SK;  
XX  
DR WPI; 1995-268290/35.  
XX  
PT New 1'-substd. spiro-indoline-3,4'-piperidine derivs. - useful as  
PT selective neurokinin-3 antagonists, e.g. for treating CNS disorders,  
PT migraine or esp. asthma.  
XX  
PS Disclosure; Column 1; 16pp; English.  
XX  
CC This sequence represents Substance P. This sequence, and those shown in  
CC AAR77311 and AAR77312 are tachykinins. These three sequences are  
CC pharmacologically active neuropeptides, and are neurokinin receptor  
CC agonists. Neurokinin receptors are widely distributed throughout the  
CC mammalian nervous system, circulatory system and peripheral tissues.  
CC Neurokinin receptors are involved in sensory perception. These  
CC sequences were used in the design and testing of neurokinin antagonists.  
CC These antagonists could be used in the treatment of conditions  
CC characterised by overstimulation of tachykinin receptors. The  
CC antagonists can also be used, for the treatment of neurodegenerative  
CC disorders (e.g. Alzheimer's disease), demyelinating diseases (e.g.

CC multiple sclerosis), respiratory diseases, ophthalmic diseases, addiction  
CC disorders, adverse immune reactions, gastrointestinal disorders, bladder  
CC function disorders, fibrosing and collagen diseases. The antagonists can  
CC also be used as diagnostic agents.

XX  
SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 16; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQOFFGLM 11

|||||

Db 1 rpkpqgffglm 11

RESULT 15

AAW33180  
ID AAW33180 standard; peptide; 11 AA.

XX  
AC AAW33180;

XX 29-JAN-1998 (first entry)

XX Mono-DTPA-Arg1 Substance P.

DE Substances P; radiolabel; diagnostic imaging; therapy;  
KW mono-DTPA-Arg1.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "DTPA-Arg"

FT Modified-site 11 /note= "amidated"

XX WO9640292-A1.

XX 19-DEC-1996.

XX 07-JUN-1996; 96WO-US09706.

XX 07-JUN-1995; 95US-0480372.

XX (MLCW ) MALLINCKRODT MEDICAL INC.

XX Srinivasan A;

XX WPI; 1997-087027/08.

XX Prepn. of pure radio-labelled peptide, e.g. for diagnostic imaging -  
PT by combining protected poly(amino:carboxylate) ligand with peptide  
PT and forming complex with radionuclide

XX Example 3; Page 12; 20pp; English.

XX Preparing a radiolabelled peptide composition, comprises combining  
CC a triamine or diamine chelating agent with a peptide, e.g. the  
CC present peptide, in a solid phase peptide synthesiser, and  
CC complexing a radionuclide with the chelate-peptide conjugate.  
CC Radiolabelled peptides or peptidomimetics can be used as diagnostic  
CC imaging agents, or in therapeutic applications, e.g. iodine(111)  
CC labelled pentatreotide can be used for somatostatin receptor  
CC imaging of neuroendocrine tumours. The radiolabelled products are  
CC obtained efficiently and inexpensively in high purity. The  
CC protected polyaminocarboxylate ligands can be added to the peptide  
CC by standard solution or solid phase peptide synthesis and  
CC deprotected with conventional reagents to give only the  
CC mono-addition product, free of di-addition product impurities. The  
CC deprotected product can be labelled with medically useful  
CC radionuclides, e.g. lanthanides or actinides, at any desired

CC location. Pre-derivatisation of individual amino acids is not  
CC required.

XX  
SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 18; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQOFFGLM 11

|||||

Db 1 rpkpqgffglm 11

RESULT 16

AAW04616  
ID AAW04616 standard; peptide; 11 AA.

XX  
AC AAW04616;

XX 13-AUG-1997 (first entry)

XX Substance P peptide for mass spectrometry analysis.

XX Mass spectrometry; polymer analysis; biopolymer analysis.

XX Synthetic.

XX WO9636986-A1.

XX 21-NOV-1996.

XX 17-MAY-1996; 96WO-US07146.

XX 19-MAY-1995; 95US-0447175.

XX 19-MAY-1995; 95US-0446055.

XX (PERS-) PERSEPTIVE BIOSYSTEMS INC.

XX Patterson DH, Tarr GE;

XX WPI; 1997-012308/01.

XX Sequencing polymers, e.g. DNA, RNA, peptide nucleic acids, proteins,  
PT etc. - by obtaining mass to charge ratios of polymer fragments,  
PT pref. using mass spectrometer, and performing statistical analysis

XX Example 2; Page 32; 86pp; English.

XX A method of obtaining sequence information about a polymer (e.g. DNA,  
CC RNA, peptide nucleic acids, proteins, peptides and carbohydrates)  
CC comprising monomers of known mass has been claimed. The present  
CC sequence represents a substance P peptide, and was used as  
CC an example as a digestion before analysis by mass spectrometry.  
CC using this novel on-plate strategy. Total sequence information  
CC from a nine well digestion can be represented in a single digestion or  
CC it is often derived from two or more wells. The methods, apparatus and  
CC kit (claimed) can be used for the analysis of polymers, particularly  
CC biopolymers, e.g. DNA, RNA, peptide nucleic acids, proteins, peptides  
CC and carbohydrates. It provides a rapid, automated and cost effective  
CC sequencing of polymers, with a statistical certainty.

XX  
SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 18; Length 11;

Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQOFFGLM 11

|||||

Db 1 rpkpqgffglm 11

## RESULT 17

AAW42973  
ID AAW42973 standard; Protein; 11 AA.XX AC  
XX DTXX AC AAW42973;  
XX DT 01-MAY-1998 (first entry)

XX DE Substrate P reporter epitope.

XX KW Beta-amyloid peptide; BAP; extracellular BAP plaque;  
KW cerebrovascular deposit; Alzheimers disease; Downs syndrome;  
KW amyloid precursor protein; APP; secretase; BAP aggregation;  
KW abnormal proteolytic cleavage; substrate P reporter epitope.

XX OS Synthetic.

XX PN US5703209-A.

XX PD 30-DEC-1997.

XX PF 05-JUN-1995; 95US-0464248.

XX PR 20-SEP-1993; 93US-0123659.

XX PR 01-MAY-1992; 92US-0877675.

XX PA (AMCY ) AMERICAN CYANAMID CO.

XX PI Jacobsen JS, Vitek MP;

XX DR WPI; 1998-076482/07.

XX PT Amyloid precursor protein fusion polypeptides - comprising APP

XX PT fragment and marker, useful for research and drug screening

XX PS Disclosure; Column 3; 84pp; English.

XX CC Peptide sequence AAW42978 represents an amyloid precursor protein (APP),  
XX CC which has a deletion of 276 amino acids to within 15 amino acids of the  
XX CC beta-amyloid peptide (BAP) domain. The protein also contains the  
XX CC Met-enkephalin reporter epitope at the carboxy terminus. Abnormal  
XX CC accumulation of extracellular BAP in plaques and cerebrovascular  
XX CC deposits is characteristic in brains of individuals suffering from  
XX CC Alzheimers disease and Downs syndrome. BAP is a poorly soluble,  
XX CC self-aggregating protein which is derived from a larger amyloid precursor  
XX CC protein (APP). APP is expressed as an integral membrane protein, and is  
XX CC cleaved by secretase, between BAP 16lys and 17leu. Cleavage at this site  
XX CC precludes amyloidogenesis and results in the release of the  
XX CC amino-terminal APP fragment. Three major isoforms of APP exist: APP-695,  
XX CC APP-751 and APP-770. These isoforms are derived by alternative splicing.  
XX CC APP-RFP 751 is constructed by ligating restriction fragments representing  
XX CC N- and C-terminal APP-751 cDNA and substrate P reporter epitope  
XX CC sequences (present sequence) APP can be used as a substrate for studying  
XX CC abnormal proteolytic cleavage which results in the release of BAP, and  
XX CC also to screen for drugs that will inhibit such cleavage.

SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 19; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11

## RESULT 18

RAY30985  
ID AAY30985 standard; peptide; 11 AA.

XX AC

XX AC AAY30985;  
XX DT 21-OCT-1999 (first entry)

XX DE Non-crosslinked protein particle peptide 34.

XX KW Non-crosslinked protein particle; diagnostic; therapy; monodisperse;  
KW albumin; haemoglobin; nanometer; micrometer; clearance.

XX OS Synthetic.

XX FH Key Location/Qualifiers  
FT Modified-site 11  
ET /note= "C-terminal amide"

XX PN US5945033-A.

XX PD 31-AUG-1999.

XX PF 12-NOV-1996; 96US-0747137.

XX PR 14-MAR-1994; 94US-0212546.

XX PR 15-JAN-1991; 91US-0641720.

XX PR 13-OCT-1992; 92US-0959560.

XX PR 01-JUN-1993; 93US-0069831.

XX PR 12-NOV-1996; 96US-0747137.

XX PA (HEMO-) HEMOSPHERE INC.

XX PI Yen RCK;

XX DR WPI; 1999-508153/42.

XX PT Non-crosslinked protein particles for therapeutic and diagnostic use

XX PS Example 22; Column 63-64; 65pp; English.

XX CC This invention describes a novel aqueous suspension of monodisperse  
XX CC particles on non-crosslinked, non-denatured albumin (50-5030 nm) which  
XX CC is stable against dissolving upon dilution with an alcohol-free aqueous  
XX CC medium. The method involves (a) forming an aqueous solution containing  
XX CC albumin and hemoglobin and (b) treating the aqueous solution with an  
XX CC alcohol to cause the solution to become turbid. The particles are useful  
XX CC as agents for in vivo administration, either of their own administration  
XX CC or as a vehicle for other therapeutic or diagnostic agents. The method  
XX CC permits the formation of albumin and hemoglobin particles in the  
XX CC nanometer and micrometer size range, in a form closer to their natural  
XX CC form than the forms of the prior art. The particles therefore constitute  
XX CC a more closely controlled agent for in vivo administration, with greater  
XX CC ease of clearance from the body after their period of usefulness.

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 20; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11

## RESULT 19

AAY03156  
ID AAY03156 standard; peptide; 11 AA.

XX AC AAY03156;

XX DT 10-JUN-1999 (first entry)

XX

DE Substance P.  
XX Opioid peptide; opioid analgesia; enhancer; opioid anaesthesia;  
KW Substance P.  
XX Synthetic.  
OS US5891842-A.  
XX 06-APR-1999.  
XX 12-APR-1996; 96US-0631434.  
XX 09-APR-1993; 93US-0044954.  
PR 12-APR-1996; 96US-0631434.  
XX (TUFTS) TUFTS COLLEGE.  
PA Kream RM;  
PI WPI; 1999-253906/21.  
XX Synergistic method for enhancing opioid analgesia and anaesthesia  
PT within a human  
XX Disclosure; Column 14; 20pp; English.  
XX This sequence represents substance P used in the method of the  
CC invention. The method is for enhancing opioid analgesia within a human  
CC subject for a duration of 15 minutes comprises concurrent administration  
CC of substance P, or one of its precursors. The method is used to elicit  
CC opioid analgesia and anaesthesia, either prior to or after the occurrence  
CC of a nociceptive event. The components have a synergistic effect. The  
CC method allows use of low doses of opioid that produce little or no  
CC physiological effect reducing conventional risks of toxicity and  
CC addiction, and allows the use of low doses of substance P and its related  
CC analogs that limit their in vivo physiological consequences. The  
CC analgesia is naloxone reversible allowing diminishment or complete  
CC elimination of opioid analgesia if desired and on demand. The treatment  
CC provides a durable analgesic effect, but only minimally disturbs and  
CC interrupts the normal metabolic processes of the body.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 100.0%; Score 61; DB 20; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQOQFFGLM 11  
DB 1 rpkpqgffglm 11

RESULT 20  
AAW92715  
ID AAW92715 standard; peptide; 11 AA.  
XX  
AC AAW92715;  
XX  
DT 30-APR-1999 (first entry)  
XX  
DE Human tachykinin agonist beta-amyloid peptide fragment #61.  
XX  
XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congenophilic angiopathy.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 10  
FT /label= MeLeu  
FT

FT  
XX US5876948-A.  
PN 02-MAR-1999.  
PD 27-JUL-1991; 91US-0737371.  
XX 29-JUL-1991; 91US-0737371.  
PR 27-JUL-1990; 90US-0559173.  
XX (CHIL-) CHILDRENS MEDICAL CENT.  
PA Yankner BA;  
XX WPI; 1999-189630/16.  
DR Screening for neurotoxin inhibitors - by testing compounds for their  
XX effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
PT Disclosure; Column 37-38; 28pp; English.  
XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis and non-inherited congenophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.  
XX  
XX Sequence 11 AA;  
SQ

Query Match 100.0%; Score 61; DB 20; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQOQFFGLM 11  
DB 1 rpkpqgffglm 11

RESULT 21  
AAW92719  
ID AAW92719 standard; peptide; 11 AA.  
XX  
AC AAW92719;  
XX  
DT 30-APR-1999 (first entry)  
XX  
DE Human tachykinin agonist beta-amyloid peptide fragment #65.  
XX  
XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congenophilic angiopathy.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 9  
FT /label= MeGly  
FT /note= "N-methyl-glycine"  
XX  
XX US5876948-A.  
XX 02-MAR-1999.  
XX 27-JUL-1991; 91US-0737371.  
XX 29-JUL-1991; 91US-0737371.  
PR 27-JUL-1990; 90US-0559173.  
XX





```
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human
XX beta-amyloid peptide fragments.
SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 20; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11
   |||||
Db 1 rpkpqffglm 11

RESULT 24
AAW92680
ID AAW92680 standard; peptide; 11 AA.
XX
AC AAW92680;
XX
DT 30-APR-1999 (first entry)
XX
DE Human tachykinin agonist beta-amyloid peptide fragment #26.
XX
KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 8 /note= "Residue is N-methyl-phenylalanine"
FT
XX
PN US5876948-A.
XX
PD 02-MAR-1999.
XX
PF 27-JUL-1991; 91US-0737371.
XX
PR 29-JUL-1991; 91US-0737371.
XX
PR 27-JUL-1990; 90US-0559173.
XX
PA (CHIL-) CHILDRENS MEDICAL CENT.
XX
PI Yankner BA;
XX
DR WPI; 1999-189630/16.
XX
PT Screening for neurotoxin inhibitors - by testing compounds for their
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells
XX
PS Disclosure; Column 21-22; 28pp; English.
XX
CC This invention describes a method for screening compounds for inhibiting
CC a neurotoxin. The method involves incubating tachykinin agonists with
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be
CC used for identifying compounds for treating diseases characterised by an
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage,
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human
XX beta-amyloid peptide fragments.
XX
SQ Sequence 11 AA;

This invention describes a method for screening compounds for inhibiting
a neurotoxin. The method involves incubating tachykinin agonists with
neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be
used for identifying compounds for treating diseases characterised by an
undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,
Down's syndrome, and the syndromes of hereditary cerebral haemorrhage,
with amyloidosis and non-inherited congophilic angiopathy with cerebral
haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human
beta-amyloid peptide fragments.

Query Match 100.0%; Score 61; DB 20; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11
   |||||
Db 1 rpkpqffglm 11

RESULT 26
AAW92676
ID AAW92676 standard; peptide; 11 AA.
XX
AC AAW92676;
XX
DT 30-APR-1999 (first entry)
XX
```

XX DE Human tachykinin agonist beta-amyloid peptide fragment #22.  
XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
XX KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX OS Homo sapiens.  
XX FH Key Location/Qualifiers  
FT Modified-site 9  
FT /label= MeGly  
FT /note= "N-methyl-glycine (Sarcosine)"  
XX PN US5876948-A.  
XX PD 02-MAR-1999.  
XX PF 27-JUL-1991; 91US-0737371.  
XX PR 29-JUL-1991; 91US-0737371.  
XX PR 27-JUL-1990; 90US-0559173.  
XX PA (CHIL-) CHILDRENS MEDICAL CENT.  
XX PI Yankner BA;  
XX XX WPI; 1999-189630/16.  
XX XX Screening for neurotoxin inhibitors - by testing compounds for their  
XX effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX PS Disclosure; Column 19-20; 28pp; English.  
XX CC This invention describes a method for screening compounds for inhibiting  
XX a neurotoxin. The method involves incubating tachykinin agonists with  
XX neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
XX used for identifying compounds for treating diseases characterised by an  
XX undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
XX Down's syndrome, and the syndromes of hereditary cerebral haemorrhage,  
XX with amyloidosis and non-inherited congophilic angiopathy with cerebral  
XX haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
XX beta-amyloid peptide fragments.  
XX SQ Sequence 11 AA;  
XX  
XX Query Match 100.0%; Score 61; DB 20; Length 11;  
XX Best Local Similarity 100.0%; Pred. No. 0.00029;  
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX Qy 1 RPKPQQFFGLM 11  
XX Db 1 rpkpqffglm 11  
XX  
XX RESULT 27  
XX AAW92731  
XX ID AAW92731 standard; peptide; 11 AA.  
XX AC AAW92731;  
XX XX 30-APR-1999 (first entry)  
XX DT Human tachykinin agonist beta-amyloid peptide fragment #77.  
XX DE Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
XX KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
XX KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX OS Homo sapiens.  
XX XX US5876948-A.  
XX PN

XX PD 02-MAR-1999.  
XX XX 27-JUL-1991; 91US-0737371.  
XX PF Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
XX KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
XX KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX OS Homo sapiens.  
XX PA (CHIL-) CHILDRENS MEDICAL CENT.  
XX PI Yankner BA;  
XX XX WPI; 1999-189630/16.  
XX XX Screening for neurotoxin inhibitors - by testing compounds for their  
XX effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX PS Disclosure; Column 43-44; 28pp; English.  
XX CC This invention describes a method for screening compounds for inhibiting  
XX a neurotoxin. The method involves incubating tachykinin agonists with  
XX neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
XX used for identifying compounds for treating diseases characterised by an  
XX undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
XX Down's syndrome, and the syndromes of hereditary cerebral haemorrhage,  
XX with amyloidosis and non-inherited congophilic angiopathy with cerebral  
XX haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
XX beta-amyloid peptide fragments.  
XX SQ Sequence 11 AA;  
XX  
XX Query Match 100.0%; Score 61; DB 20; Length 11;  
XX Best Local Similarity 100.0%; Pred. No. 0.00029;  
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX Qy 1 RPKPQQFFGLM 11  
XX Db 1 rpkpqffglm 11  
XX  
XX RESULT 28  
XX AAW79662  
XX ID AAW79662 standard; peptide; 11 AA.  
XX AC AAW79662;  
XX XX 02-MAR-1999 (first entry)  
XX DT Substance P derivative having complex glycosylation.  
XX DE Substance P; mannose; glycosylation; solubility.  
XX KW Synthetic.  
XX OS Key Location/Qualifiers  
XX FH Region 1..4  
XX FT /note= "optionally the first four residues may be  
XX FT deleted, leaving SP(5-11)"  
XX FT Modified-site 5  
XX FT /note= "the side chain amide group is N-substituted  
XX FT with N-acetyl-D-glucosamine (GlcNAc) which in turn  
XX FT is extended in the 4-position with a complex type  
XX FT sugar chain, a high mannose type sugar chain or a  
XX FT mixed type sugar chain"  
XX FT Modified-site 11  
XX FT /note= "Met-NH2, i.e. C-terminal amide"  
XX XX JPL0306099-A.  
XX PN 17-NOV-1998.  
XX PD 28-NOV-1997; 97JP-0343979.  
XX PF

XX  
PR 04-MAR-1997; 97JP-0065372.  
XX  
PA (NOGK ) ZH NOGUCHI KENKYUSHO.  
XX  
DR WPI; 1999-054306/05.  
XX  
XX New substance P derivatives with side chain containing sugar - has  
PT improved solubility  
XX  
XX Claim 1; Page 2; 8pp; Japanese.  
XX  
CC The sequence represents the peptide portion of a new Substance P  
CC derivative having complex glycosylation on the Gln(5) position. The  
CC derivative has improved solubility compared with Substance P.  
XX  
SQ Sequence 11 AA;  
  
Query Match 100.0%; Score 61; DB 20; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RPKPQQFFGLM 11  
Db 1 rpkpqgffglm 11  
  
RESULT 29  
AAW79663  
ID AAW79663 standard; peptide; 11 AA.  
XX  
AC AAW79663;  
XX  
DT 02-MAR-1999 (first entry)  
XX  
DE Substance P derivative having complex glycosylation.  
XX  
KW Substance P; mannose; glycosylation; solubility.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 6  
FT /note= "the side chain amide group is N-substituted  
FT with N-acetyl-D-glucosamine (GlcNAc) which in turn  
FT is extended in the 4-position with a complex type  
FT sugar chain, a high mannose type sugar chain or a  
FT mixed type sugar chain"  
FT Modified-site 11  
FT /note= "Met-NH2, i.e. C-terminal amide"  
XX  
PN JP10306099-A.  
XX  
PD 17-NOV-1998.  
XX  
PF 28-NOV-1997; 97JP-0343979.  
XX  
PR 04-MAR-1997; 97JP-0065372.  
XX  
PA (NOGK ) ZH NOGUCHI KENKYUSHO.  
XX  
DR WPI; 1999-054306/05.  
XX  
XX New substance P derivatives with side chain containing sugar - has  
PT improved solubility  
XX  
XX Claim 1; Page 2; 8pp; Japanese.  
XX  
CC The sequence represents the peptide portion of a new Substance P  
CC derivative having complex glycosylation on the Gln(6) position. The  
CC derivative has improved solubility compared with Substance P.  
XX

SQ Sequence 11 AA;  
  
Query Match 100.0%; Score 61; DB 20; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RPKPQQFFGLM 11  
Db 1 rpkpqgffglm 11  
  
RESULT 30  
AAB18483  
ID AAB18483 standard; peptide; 11 AA.  
XX  
AC AAB18483;  
XX  
DT 15-JAN-2001 (first entry)  
XX  
DE Peptide substrate used to test prolyl-tripeptidyl peptidase activity.  
XX  
KW Prolyl tripeptidyl-peptidase; amidolytic activity; periodontal disease;  
KW gingivitis; periodontitis.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT /note= "hydrogen attached"  
FT Modified-site 11  
FT /note= "amidated residue"  
XX  
PN WO200052147-A2.  
XX  
PD 08-SEP-2000.  
XX  
PF 03-MAR-2000; 2000WO-US05551.  
XX  
PR 05-MAR-1999; 99US-0123148.  
XX  
XX (UYGE-) UNIV GEORGIA RES FOUND INC.  
PA (TRAV/) TRAVIS J.  
PA (POTE/) POTEMPA J.  
PA (BANB/) BANBULA A.  
XX  
PI Travis J, Potempa J, Banbula A;  
XX WPI; 2000-594181/56.  
XX  
PT Prolyl tripeptidyl-peptidase, active analog, fragment or variant useful  
PT for identifying its inhibitor which is useful for protecting an animal  
PT from a periodontal disease such as gingivitis and periodontitis -  
XX  
PS Example 4; Page 37; 58pp; English.  
XX  
CC The present sequence represents a substrate which was used to test  
CC the activity of prolyl tripeptidyl-peptidases PRP-A and PRP IV. The  
CC prolyl tripeptidyl-peptidase has an amidolytic activity, and cleaves  
CC a peptide bond in a target polypeptide having at least 4 amino acids.  
CC This bond is between a proline and an amino acid attached to the  
CC alpha-carboxyl group end of the proline. The polypeptide is useful for  
CC identifying inhibitors. These inhibitors are then useful for reducing  
CC the growth of bacterium or for protecting an animal from a periodontal  
CC disease such as gingivitis and periodontitis caused by Porphyromonas  
CC gingivalis.  
XX  
SQ Sequence 11 AA;  
  
Query Match 100.0%; Score 61; DB 21; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RPKPQQFFGLM 11  
|||||  
Db 1 rpkpqffglm 11

## RESULT 31

AAB23027  
ID AAB23027 standard; peptide; 11 AA.

XX AC AAB23027;  
XX 16-JAN-2001 (first entry)  
XX Human/rat tachykinin Substance P.  
XX Substance P; tachykinin; human; rat; magnesium binding defect;  
KW sodium sensitive essential hypertension; insulin resistance;  
KW type 2 diabetes; antibody; immunoassay; quantification.  
XX Homo sapiens.  
OS Rattus sp.

XX FH Key Location/Qualifiers  
FT Modified-site 11  
FT /note= "C-terminal amide"

XX PN WO2000054053-A1.

XX PD 14-SEP-2000.

XX PF 09-MAR-2000; 2000WO-US03707.

XX PR 10-MAR-1999; 99US-0265690.

XX PA (WELL)/ WELLS I C.

XX PI Wells IC;

XX DR WPI; 2000-587457/55.

XX Detecting magnesium binding defects associated with abnormal  
PT physiological states such as sodium-sensitive essential hypertension  
PT and type 2 insulin-resistant diabetes mellitus, comprises measuring a  
PT specific pentapeptide in blood -

XX PS Disclosure; Page 5; 21pp; English.

XX The invention relates to a method for detecting magnesium binding  
CC defects. The method comprises quantitating a tachykinin C-terminal  
CC pentapeptide (e.g., AAB23025) and its degradation products (e.g.,  
CC AAB23026) in blood using an antibody specific for the generalised  
CC mammalian tachykinin C-terminal pentapeptide  
CC Phe-(Phe/Val)-Gly-Leu-Met-NH<sub>2</sub> (AAB23028). The method is useful for  
CC detecting cellular magnesium binding defects which are associated with  
CC abnormal physiological states such as sodium-sensitive essential  
CC hypertension and type 2 diabetes mellitus. The present sequence  
CC represents the tachykinin Substance P from human and rat. C-terminal  
CC fragments (AAB23025, AAB23026) of the present sequence may be assayed  
CC according to the method of the invention.

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 21; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RPKPQQFFGLM 11

Db 1 rpkpqffglm 11

## RESULT 32

AAY32382  
ID AAY32382 standard; Peptide; 11 AA.

XX AC AAY32382;

XX 28-FEB-2000 (first entry)

XX Cell differentiation, proliferation and maintenance factor peptide.

XX Cell differentiation; cell proliferation; cell maintenance;  
KW ectoderm-like cell; embryonic stem cell; pluripotent cell;  
KW gene therapy; cell therapy; tissue transplant; organ transplant;  
KW xerograft; allotransplant; concomitant transplantation;  
KW transgenic animal; substance P.

XX OS Synthetic.

XX PN WO9953021-A1.

XX PD 21-OCT-1999.

XX PF 09-APR-1999; 99WO-AU00265.

XX PR 09-APR-1998; 98AU-0002912.

XX PR 23-SEP-1998; 98AU-0006097.

XX PA (BRES-) BRESAGEN LTD.

XX PI Bettess MD, Rathjen PD, Rathjen J;

XX DR WPI; 2000-061970/05.

XX New isolated biologically active factor capable of influencing  
PT differentiation, proliferation or maintenance of pluripotent cells

XX Claim 3; Page 123; 189pp; English.

XX This sequence represents a peptide (substance P free acid); that can  
CC form the low mol.wt. component of a novel biologically active factor  
CC that is capable of influencing the differentiation, proliferation  
CC and/or maintenance of pluripotent cells. The factor consists of a  
CC low mol.wt. component selected from Pro, Pro-Ala, Ala-Pro-Gly,  
CC Pro-OH-Pro, Gly-Pro-Ala, Gly-Pro-OH-Pro, a peptide given in  
CC AAY32378-82, or a protease digested (including collagenase digested)  
CC collagen fragment, and a high mol.wt. component such as fibronectin.  
CC The biologically active factor is obtained from conditioned media of  
CC hepatic or hepatoma cells or cell lines or extraembryonic endodermal  
CC cells or cell lines. The factor is capable of causing the  
CC transition of pluripotent cells (e.g. embryonic stem cells in  
CC adherent culture and in suspension culture) to pluripotent cells  
CC having different properties, more specifically primitive  
CC ectoderm-like (EPL) cells. The factor is also capable of  
CC maintaining and supporting proliferation of these cells in vitro.  
CC It also allows the isolation and maintenance of EPL cells derived  
CC from in vitro and in vivo primitive ectoderm. These cells can be  
CC used in allo-, concomitant- or xeno-transplantation, cell therapy,  
CC tissue and organ augmentation or replacement, and gene therapy.  
CC They can also be used for producing chimeric or transgenic animals.

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 21; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RPKPQQFFGLM 11

Db 1 rpkpqffglm 11

## RESULT 33

AAG62768  
ID AAG62768 standard; peptide; 11 AA.  
XX  
AC AAG62768;  
XX  
DT 17-SEP-2001 (first entry)  
XX  
DE Amino acid sequence of substance P.  
XX  
KW Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 11  
FT /note= "amidated residue"  
XX  
PN WO200153336-A1.  
XX  
PD 26-JUL-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01529.  
XX  
PR 19-JAN-2000; 2000US-0489667.  
XX  
PA (ALLR ) ALLERGAN SALES INC.  
XX  
PI Donovan S;  
XX  
DR WPI; 2001-451900/48.  
XX  
XX Agent useful for treating pain comprises a clostridial neurotoxin (or component) attached to a targeting moiety  
XX  
PS Disclosure; Page 61; 77pp; English.  
XX  
CC The specification describes an agent, comprising a clostridial neurotoxin attached to a targeting moiety, where the targeting moiety is selected from transmission compounds, and compounds substantially similar to the transmission compounds. The agent may be used for treating pain, where the clostridial neurotoxin component is derived from botulinum toxin selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
CC The targeting moiety comprises a light chain and an amine end segment of a heavy chain and comprises substance P as the targeting moiety. The pain alleviating effects persist for 2-6 months. The present sequence represents substance P, and is used in the course of the invention.  
XX  
SQ Sequence 11 AA;  
XX  
Query Match 100.0%; Score 61; DB 22; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQOFFGLM 11  
Db 1 rpkpqgffglm 11  
RESULT 34  
AAG99354  
ID AAG99354 standard; Peptide; 11 AA.  
XX  
AC AAG99354;  
XX  
DT 25-SEP-2001 (first entry)  
XX  
DE Substance P peptide.  
XX  
KW Atypical tachykinin; ATT; human; hypertension.  
XX  
OS Unidentified.  
XX  
PN WO200146415-A1.  
XX  
PD 28-JUN-2001.  
XX  
PF 21-DEC-2000; 2000WO-JP09083.  
XX  
PR 21-DEC-1999; 99JP-0362638.  
XX  
PR 10-MAR-2000; 2000JP-0066714.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Itoh Y, Nishi K, Kitada C, Inatomi N;  
XX  
PI WPI; 2001-441676/47.  
XX  
DR Atypical tachykinin peptides of human origin and DNA encoding them for screening potential agents for treatment of hypertension  
XX  
PT  
XX  
PS Disclosure; Page 9; 153pp; Japanese.  
XX  
CC The present invention relates to atypical tachykinin proteins of human origin and their esters, amides, salts and partial peptides. These can be used in the treatment, prevention and diagnosis of hypertension. The present sequence is a protein fragment described in the exemplification of the invention.  
XX  
SQ Sequence 11 AA;  
XX  
Query Match 100.0%; Score 61; DB 22; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQOFFGLM 11  
Db 1 rpkpqgffglm 11  
RESULT 35  
AAB84527  
ID AAB84527 standard; peptide; 11 AA.  
XX  
AC AAB84527;  
XX  
DT 05-SEP-2001 (first entry)  
XX  
DE Amino acid sequence of human substance P.  
XX  
KW Substance P; cell toxin; Pseudomonas exotoxin; cell ablation;  
KW NK-1 receptor; chronic pain; tumour; neurological dysfunction;  
KW basal ganglia; cholinergic interneuron; Parkinson's disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200131020-A1.  
XX  
PD 03-MAY-2001.  
XX  
PF 20-OCT-2000; 2000WO-US29064.  
XX  
PR 22-OCT-1999; 99US-0161159.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
PI Fitzgerald DJ, Iadarola MJ;  
XX  
PI WPI; 2001-417560/44.  
XX  
PT Making cell toxin to treat chronic pain, by forming substance P-pseudomonas exotoxin disulfide-linked conjugate, by reacting modified exotoxin and substance P having additional cysteine residue at its N-terminus  
XX

PS Disclosure; Page 10; 54pp; English.

CC The present sequence represents a human substance P. The peptide is  
CC used to produce a cell toxin. The cell toxin comprises a substance  
CC P-pseudomonas exotoxin disulfide-linked conjugate. The cell toxin is  
CC useful for ablating NK-1 receptor expressing cells, such as dorsal horn  
CC cell, a stratum cell or a brain parenchyma cell, for treating chronic  
CC pain in epineurium cells, perineurium cells, nerve ganglia, nerve  
CC sheaths, nerve linings, meninges, planar cells, arachnoid membrane  
CC cells, duramembrane cells, cells lining a joint or brain or spinal cord  
CC parenchymal cells, without significantly affecting basal nociceptive  
CC responses. The cell toxin is thus useful for treating chronic pain or  
CC tumours that binds substance P. It is also useful for neurological  
CC dysfunctions of the basal ganglia by targeting cholinergic interneurons  
CC that express substance P e.g. Parkinson's disease.

XX Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11  
|||||  
Db 1 rpkpqffgflm 11

RESULT 36

AAB98866  
ID AAB98866 standard; Peptide; 11 AA.

AC AAB98866;

DT 14-AUG-2001 (first entry)

DE Chimeric analgesic peptide #22.

KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.

OS Synthetic.

FH Key Location/Qualifiers  
FT Modified-site 11

FT /label= OTHER  
FT /note= "C-terminal amide"

PN WO200130371-A2.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-US29789.

XX 28-OCT-1999; 99US-0428692.

XX (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;

DR WPI; 2001-397593/42.

XX New chimeric peptides used for treating pain comprise opioid receptor  
PT binding group and nociceptive receptor binding group

XX Claim 10; Page 15; 34pp; English.

CC The present invention describes a number of chimeric peptides comprising  
CC an opioid receptor binding moiety and a nociceptive receptor binding  
CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
CC opioid receptor based peptides alone, tolerance does not result from  
CC their long-term use. The present sequence is one of the peptides of the  
CC invention.

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11  
|||||  
Db 1 rpkpqffgflm 11

RESULT 37

AAB82070  
ID AAB82070 standard; peptide; 11 AA.

XX AAB82070;

XX 22-JUN-2001 (first entry)

XX Substance P.

XX Antigen; immunostimulant; vaccine; pharmaceutical composition; antiviral;  
KW viral infection; substance P.

OS Unidentified.

FH Key Location/Qualifiers  
FT Modified-site 11

FT /note= "C-terminal amide"

XX WO200124822-A2.

XX 12-APR-2001.

XX 02-OCT-2000; 2000WO-EP09657.

XX 01-OCT-1999; 99AT-0001680.

XX (CIST-) CISTEM BIOTECHNOLOGIES GMBH.

XX Fleitmann J, Mattner F, Buschle M, Melling J;

XX WPI; 2001-290577/30.

XX New pharmaceutical composition comprising an antigen, an  
PT immunostimulating substance and a polycationic polymer, useful in  
PT manufacturing vaccines

XX Example 3; Page 14; 20pp; English.

CC The present invention relates to a pharmaceutical composition comprising  
CC (a) an antigen; (b) an immunostimulating substance consisting of  
CC neuroactive compounds, hormones, compounds having growth hormone activity  
CC or their mixtures; and (c) a polycationic polymer. The composition is  
CC useful in manufacturing vaccines. To illustrate the present invention, a  
CC murine tyrosinase related protein-2 peptide (TRP-2 peptide; see  
CC AAB82064), was used. Mice were injected subcutaneously with either the  
CC TRP-2 peptide, TRP-2 peptide + poly-L-arginine 60 (pR60) or TRP-2 peptide  
CC + pR60 + substance P (the present peptide). Animals were sacrificed 10  
CC days post injection, and spleen tissue was harvested. Lymphocytes were  
CC prepared from the spleen tissue and were re-stimulated with TRP-2 peptide  
CC or with an ovalbumin-derived peptide (AAB82065), with the same major  
CC histocompatibility complex (MHC) restriction serving as negative control.  
CC Spots representing single T cells specific for the peptide used for  
CC re-stimulation were counted. No spots were detected when the ovalbumin  
CC derived peptide was used, while TRP-2 peptide + pR60 + substance P showed  
CC the highest number of spots or single T cells.

XX SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
| | | | | | | | | |  
Db 1 rpKpqgffglm 11

RESULT 38  
AAB91436  
ID AAB91436 standard; Peptide; 11 AA.

XX AC AAB91436;

DT 22-JUN-2001 (first entry)

DE Tachykinins peptide SEQ ID NO:612.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.  
OS Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US13576.

XX 17-MAY-1999; 99US-0134406.

XX 10-SEP-1999; 99US-0153406.

XX 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT

PS Disclosure; Page 399; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.

XX Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
| | | | | | | | | |  
Db 1 rpKpqgffglm 11

RESULT 39  
AAB91449  
ID AAB91449 standard; Peptide; 11 AA.

XX AC AAB91449;

DT 22-JUN-2001 (first entry)

DE Tachykinins peptide SEQ ID NO:625.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.  
OS Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US13576.

XX 17-MAY-1999; 99US-0134406.

XX 10-SEP-1999; 99US-0153406.

XX 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT

PS Disclosure; Page 403; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.

XX Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
| | | | | | | | | |  
Db 1 rpKpqgffglm 11

```
RESULT 40
AAB91450
ID AAB91450 standard; Peptide; 11 AA.
XX
AC AAB91450;
XX
DE 22-JUN-2001 (first entry)
XX
DT Tachykinins peptide SEQ ID NO:626.
XX
DE Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
XX
PR 10-SEP-1999; 99US-0153406.
XX
PR 15-OCT-1999; 99US-0159783.
XX
XX (CONJ-) CONJUCHEM INC.
XX
PA Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thiбаudeau K;
XX
PI WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT
XX
PS Disclosure; Page 403; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11
Db 1 rpkpqffgflm 11

RESULT 41
AAB50544
ID AAB50544 standard; Peptide; 11 AA.
```

```
XX
AC AAB50544;
XX
DT 16-MAR-2001 (first entry)
XX
DE Prolyl endopeptidase inhibitor substance P peptide.
XX
KW Prolyl endopeptidase inhibitor; PEP inhibitor; central nervous system;
KW CNS; nootropic; brain function disorder; Alzheimer's disease; amnesia.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
FH Modified-site 11
FT /note= "amidated"
XX
XX WO200071144-A1.
XX
XX 30-NOV-2000.
XX
XX 16-MAY-2000; 2000WO-JP03135.
XX
XX 19-MAY-1999; 99JP-0138791.
XX
XX (DOME-) DOMER INC.
XX
XX Kayahara H, Tsukahara K, Inagaki T;
XX
XX WPI; 2001-070833/08.
XX
XX Prolyl endopeptidase inhibitor comprises cereal extract including new
XX ketone compound.
XX
XX Disclosure; Fig 1; 27pp; Japanese.
XX
CC The present invention describes prolyl endopeptidase (PEP) inhibitors
CC comprising a cereal extract. Also described are:
CC (i) a 7-octadecenyl-7,10-henecosadienyl ketone;
CC (ii) germinating brown rice having prolyl endopeptidase inhibitory
CC activity for preventing and/or relieving brain function disorders; and
CC (iii) foods for preventing or relieving brain function disorders
CC comprising the above PEP inhibitor or the above germinated brown rice.
CC The PEP inhibitors can have central nervous system (CNS) and nootropic
CC activity. The PEP inhibitors can be used for preventing and relieving
CC brain function disorders including Alzheimer's disease and amnesia.
CC The present sequence represents a PEP inhibitor peptide given in the
CC exemplification of the present invention.
XX
SQ Sequence 11 AA;

Query Match 100.0%; Score 61; DB 22; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11
Db 1 rpkpqffgflm 11

RESULT 42
AAB50306
ID AAB50306 standard; peptide; 11 AA.
XX
AC AAB50306;
XX
XX 08-MAR-2001 (first entry)
XX
XX Substance P.
XX
XX Antibacterial; Botulinum toxin inhibitor; BttxB;
KW previn; tetanus neurotoxin; buforinlin; substance P.
XX
XX
```



OS Unidentified.  
PN WO200069891-A2.  
XX  
XX 23-NOV-2000.  
PD  
XX  
PF 15-MAY-2000; 2000WO-US13215.  
XX  
PR 17-MAY-1999; 99US-0134446.  
XX  
PA (USSA ) US DEPT OF THE ARMY.  
XX  
XX Gordon RK, Moorad DR, Doctor BP, Garcia GE;  
PI WPI; 2001-025001/03.  
XX  
DR Novel Previn compounds useful for inhibiting the protease activity of  
PT Botulinum B and tetanus toxins -  
XX  
XX Claim 7; Page 29; 47pp; English.  
XX  
XX The present sequence was investigated in the search for Botulinum  
CC toxin inhibitors (BttxB). Previn compounds which inhibit the enzymatic  
CC activity of BttxB and tetanus neurotoxins were isolated. Previn  
CC may be used to construct compounds such as buforinins.  
XX  
XX Sequence 11 AA;  
SQ  
Query Match 100.0%; Score 61; DB 22; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQOQFFGLM 11  
Db 1 rpkpqgffglm 11  
RESULT 43  
AAR32798  
ID AAR32798 standard; peptide; 12 AA.  
XX  
AC AAR32798;  
XX  
DT 17-JUN-1993 (first entry)  
XX  
DE Tyr-1 substance P used for binding assay.  
XX  
XX human substance P receptor protein; SP; neurotransmitter;  
KW neuromodulator; central nervous system; peripheral nervous system;  
KW gastrointestinal disorders; inflammation; immune disease.  
XX  
OS Homo sapiens.  
XX  
XX WO9303137-A.  
PN  
XX  
PD 18-FEB-1993.  
XX  
XX 05-AUG-1992; 92WO-US05532.  
PF  
XX  
PR 07-AUG-1991; 91US-0741200.  
XX  
XX (UNIW ) UNIV WASHINGTON.  
PA  
XX  
PI Krause JE;  
XX  
XX WPI; 1993-076495/09.  
DR  
XX  
XX New human substance P receptor protein and DNA encoding it - used  
PT e.g. for screening substance P antagonists  
XX  
XX Example; Page 8; 40pp; English.  
PS  
XX

CC This sequence represents Tyr-1 substance P and was used in its  
CC 125-Iodinated form in a ligand binding assay of COS-7 cells  
CC transfected with substance P receptor coding plasmids (see AAQ37210).  
XX  
XX Sequence 12 AA;  
SQ  
Query Match 100.0%; Score 61; DB 14; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQOQFFGLM 11  
Db 2 rpkpqgffglm 12  
RESULT 44  
AAR85244  
ID AAR85244 standard; peptide; 12 AA.  
XX  
AC AAR85244;  
XX  
DT 18-AUG-1997 (first entry)  
XX  
DE Substance P analogue peptide Cys-SP.  
XX  
XX Ligand; antibody; receptor; SELEX; random library; amplification; PCR;  
KW Systematic Evolution of Ligands by Exponential enrichment; primer;  
KW polymerase chain reaction; structure; mimicry; substance P; tachykinin;  
KW neuropeptide; rheumatoid arthritis; atherosclerosis; cancer;  
KW diabetic retinopathy.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 1  
FT /note= "Ac-Arg"  
XX  
XX WO9530775-A1.  
PN  
XX  
PD 16-NOV-1995.  
XX  
XX 03-MAY-1995; 95WO-US05600.  
XX  
XX 21-DEC-1994; 94US-0361795.  
PR  
XX 06-MAY-1994; 94US-0238863.  
PR  
XX 24-MAY-1994; 94US-0248632.  
PR  
XX 09-SEP-1994; 94US-0303362.  
PR  
XX 11-JUN-1990; 90US-0536428.  
PR  
XX 10-JUN-1991; 91US-0714131.  
PR  
XX 21-OCT-1992; 92US-0964624.  
XX  
PA (UYRE-) UNIV RES CORP.  
XX  
XX Allen P, Doudna JA, Feigon J, Gold L, Nieuwlandt D;  
PI Schneider DJ, Sullenger BA, Wecker M;  
XX  
XX WPI; 1995-404132/51.  
DR  
XX  
XX Systematic evolution of ligands by exponential enrichment - for  
PT identifying nucleic acid ligands used in the treatment of, e.g. type  
PT B insulin resistance and HIV  
XX  
XX Example 11; Fig 8; 209pp; English.  
PS  
XX  
XX The invention relates to a novel method of isolating ligands that bind  
CC to target proteins e.g. antibodies or receptors, which bind other  
CC proteins or ligands. The method, designated Systematic Evolution of  
CC Ligands by Exponential enrichment (SELEX), comprises generating a library  
CC of random oligonucleotide sequences, about 40-60 nucleotides in length,  
CC and binding these sequences to the target proteins. After removal of  
CC unbound material, the remaining bound nucleotide sequences are amplified  
CC e.g. by PCR, and the newly amplified material is bound again with the

CC target protein. This cycle continues until a sufficiently pure  
CC oligonucleotide sequence is isolated. The method allows the isolation of  
CC oligonucleotide sequences which structurally mimic the target protein's  
CC ligand. This peptide represents an analogue of Substance P (AAR85243) in  
CC which the N-terminal amine has been acylated in order to determine  
CC whether this functional group interacts with nucleic acid ligands binding  
CC substance P (see AAT06098-130). The ligands can be used to block the  
CC activity of Substance P and is useful in the treatment of e.g. rheumatoid  
CC arthritis, atherosclerosis, diabetic retinopathy or cancer.

XX Sequence 12 AA;

Query Match 100.0%; Score 61; DB 16; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RPKPQQFFGLM 11  
| | | | | | | | | |  
Db 1 rpkpqffglm 11

#### RESULT 45

AA03157  
ID AAY03157 standard; peptide; 12 AA.

XX AC AAY03157;

XX DT 10-JUN-1999 (first entry)

XX DE Substance P-Glycine.

XX KW Opioid peptide; opioid analgesia; enhancer; opioid anaesthesia;  
XX substance P.

XX OS Synthetic.

XX PN US5891842-A.

XX PD 06-APR-1999.

XX PF 12-APR-1996; 96US-0631434.

XX PR 09-APR-1993; 93US-0044954.

XX PR 12-APR-1996; 96US-0631434.

XX PA (TUFT ) TUFTS COLLEGE.

XX PI Cream RM;

XX DR WPI; 1999-253906/21.

XX PT Synergistic method for enhancing opioid analgesia and anaesthesia  
XX within a human

XX PS Disclosure; Column 14; 20pp; English.

XX This sequence represents substance P used in the method of the  
CC invention. The method is for enhancing opioid analgesia within a human  
CC subject for a duration of 15 minutes comprises concurrent administration  
CC of substance P, or one of its precursors. The method is used to elicit  
CC opioid analgesia and anaesthesia, either prior to or after the occurrence  
CC of a nociceptive event. The components have a synergistic effect. The  
CC method allows use of low doses of opioid that produce little or no  
CC physiological effect reducing conventional risks of toxicity and  
CC addiction, and allows the use of low doses of substance P and its related  
CC analogs that limit their in vivo physiological consequences. The  
CC analgesia is naloxone reversible allowing diminishment or complete  
CC elimination of opioid analgesia if desired and on demand. The treatment  
CC provides a durable analgesic effect, but only minimally disturbs and  
CC interrupts the normal metabolic processes of the body.

XX Sequence 12 AA;

Query Match 100.0%; Score 61; DB 20; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RPKPQQFFGLM 11  
| | | | | | | | | |  
Db 1 rpkpqffglm 11

#### RESULT 46

AAW94412

ID AAW94412 standard; peptide; 12 AA.

XX AC AAW94412;

XX DT 15-APR-1999 (first entry)

XX DE Cancer protease-sensitive amino acid linker pAP-215 and pAP-216.

XX KW Ricin-like toxin; cancer; viral infection; parasitic infection;  
XX linker; B chain; A chain; protease; fungal infection; malaria;  
XX leucocyte proliferation; cytomegalovirus; herpes; hepatitis;  
XX rhinovirus; laryngeotracheitis; poliomyelitis; varicella zoster;  
XX cystic fibrosis; multiple sclerosis.

XX OS Unidentified.

XX OS Synthetic.

XX PN W09849311-A2.

XX PD 05-NOV-1998.

XX PF 30-APR-1998; 98WO-CA00394.

XX PR 29-OCT-1997; 97US-0063715.

XX PR 30-APR-1997; 97US-0045148.

XX PA (DNOV-) DE NOVO ENZYME CORP.

XX PI Borgford T;

XX DR WPI; 1999-009431/01.

XX PT New nucleic acid encoding ricin-like toxin with an interchain linker  
XX cleaved by protease - is specific for diseased cells, useful for,  
XX e.g. killing selectively cancer or infected cells

XX PS Claim 24; Fig 21; 352pp; English.

XX The present invention describes new purified and isolated nucleic acids  
CC (i) encoding: (1) the A and B chains of a ricin-like toxin (II); and  
CC (ii) a heterologous linker, joining the two chains and including a  
CC cleavage recognition site for a disease-specific protease (III). Also  
CC described are: (1) plasmids or baculovirus transfer vectors that contain  
CC (1); and (2) recombinant protein (IV) consisting of the A and B chains  
CC of (II) joined by the specified linker. (IV), produced by expression of  
CC (I) in host cells, are used to inhibit or kill diseased cells that  
CC produce (III), particularly for treating cancers (e.g. leucocyte  
CC proliferation; cancer of ovary, pancreas, breast or prostate; glioma) or  
CC infections caused by fungi, parasites (e.g. malaria) or viruses (e.g.  
CC cytomegalovirus (CMV), herpes, hepatitis, rhinovirus, laryngeotracheitis,  
CC poliomyelitis or varicella zoster), also cystic fibrosis and multiple  
CC sclerosis. Alternatively, (I) is used to express (IV) in vivo. (IV) is  
CC toxic specifically for (III)-expressing cells and does not depend for  
CC specificity on a cell-binding component. When used to treat virus-  
CC infected cells, transcytosis and cytotoxicity of (IV) are increased by  
CC retrograde translocation from endoplasmic reticulum to cytoplasm (which  
CC some viruses exploit to avoid immune detection), so selectivity and  
CC safety are further improved. (IV) are not toxic until chain A is  
CC released and this occurs only in target cells. The present sequence  
CC represents a specifically claimed cancer protease-sensitive amino acid

CC linker from the present invention.

XX Sequence 12 AA;

Query Match 100.0%; Score 61; DB 20; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
Db 1 rpkpqgffglm 11

RESULT 47

AAG62769

ID AAG62769 standard; peptide; 12 AA.

XX AC AAG62769;

XX DT 17-SEP-2001 (first entry)

DE Amino acid sequence of substance P precursor.

XX Clostridial neurotoxin; pain; botulinum toxin; Substance P.

XX Unidentified.

XX WO200153336-A1.

XX PD 26-JUL-2001.

XX PF 17-JAN-2001; 2001WO-US01529.

XX PR 19-JAN-2000; 2000US-0489667.

XX PA (ALLR ) ALLERGAN SALES INC.

XX PI Donovan S;

XX WPI; 2001-451900/48.

XX Agent useful for treating pain comprises a clostridial neurotoxin (or component) attached to a targeting moiety

PS Disclosure; Page 62; 77pp; English.

XX The specification describes an agent, comprising a clostridial neurotoxin attached to a targeting moiety, where the targeting moiety is selected from transmission compounds, and compounds substantially similar to the transmission compounds. The agent may be used for treating pain, where the clostridial neurotoxin component is derived from botulinum toxin selected from botulinum types A, B, C, D, E, F, G and mixtures of these. The targeting moiety comprises a light chain and an amine end segment of a heavy chain and comprises Substance P as the targeting moiety. The pain alleviating effects persist for 2-6 months. The present sequence represents substance P precursor, and is used in the course of the invention.

XX Sequence 12 AA;

Query Match 100.0%; Score 61; DB 22; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
Db 1 rpkpqgffglm 11

RESULT 48

AAG62772

ID AAG62772 standard; peptide; 12 AA.

XX AC AAG62772;

XX DT 17-SEP-2001 (first entry)

DE Amino acid sequence of carboxy-ester substance P precursor.

XX Clostridial neurotoxin; pain; botulinum toxin; Substance P.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 12  
ET /note= "methylated residue"

XX WO200153336-A1.

XX PD 26-JUL-2001.

XX PF 17-JAN-2001; 2001WO-US01529.

XX PR 19-JAN-2000; 2000US-0489667.

XX PA (ALLR ) ALLERGAN SALES INC.

XX PI Donovan S;

XX WPI; 2001-451900/48.

XX Agent useful for treating pain comprises a clostridial neurotoxin (or component) attached to a targeting moiety

PS Disclosure; Page 64; 77pp; English.

XX The specification describes an agent, comprising a clostridial neurotoxin attached to a targeting moiety, where the targeting moiety is selected from transmission compounds, and compounds substantially similar to the transmission compounds. The agent may be used for treating pain, where the clostridial neurotoxin component is derived from botulinum toxin selected from botulinum types A, B, C, D, E, F, G and mixtures of these. The targeting moiety comprises a light chain and an amine end segment of a heavy chain and comprises Substance P as the targeting moiety. The pain alleviating effects persist for 2-6 months. The present sequence represents a substance P precursor, and is used in the course of the invention.

XX Sequence 12 AA;

Query Match 100.0%; Score 61; DB 22; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
Db 1 rpkpqgffglm 11

RESULT 49

AAG62775

ID AAG62775 standard; peptide; 12 AA.

XX AC AAG62775;

XX DT 17-SEP-2001 (first entry)

DE Amino acid sequence of carboxy-ester substance P precursor.

XX Clostridial neurotoxin; pain; botulinum toxin; Substance P.

XX Synthetic.

```
FH Key Location/Qualifiers
FT Modified-site 12
FT /note= "ethylated residue"
XX
XX WO200153336-A1.
XX
XX 26-JUL-2001.
XX
XX 17-JAN-2001; 2001WO-US01529.
XX
XX 19-JAN-2000; 2000US-0489667.
XX
XX (ALLR ) ALLERGAN SALES INC.
XX
XX Donovan S;
XX
XX WPI; 2001-451900/48.
XX
XX Agent useful for treating pain comprises a clostridial neurotoxin (or
XX component) attached to a targeting moiety .
XX
XX Disclosure; Page 67; 77pp; English.
XX
XX The specification describes an agent, comprising a clostridial neurotoxin
XX attached to a targeting moiety, where the targeting moiety is selected
XX from transmission compounds, and compounds substantially similar to the
XX transmission compounds. The agent may be used for treating pain, where
XX the clostridial neurotoxin component is derived from botulinum toxin
XX selected from botulinum types A, B, C, D, E, F, G and mixtures of these.
XX The targeting moiety comprises a light chain and an amine end segment of
XX a heavy chain and comprises Substance P as the targeting moiety. The pain
XX alleviating effects persist for 2-6 months. The present sequence
XX represents a substance P precursor, and is used in the course of the
XX invention.
XX
XX Sequence 12 AA;

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQQFFGLM 11
Db 1 rpqpqqffglm 11

RESULT 50
AAB84528
ID AAB84528 standard; peptide; 12 AA.
XX
XX AAB84528;
XX
XX 05-SEP-2001 (first entry)
XX
XX Amino acid sequence of a modified substance P.
XX
XX Substance P; cell toxin; Pseudomonas exotoxin; cell ablation;
XX NK-1 receptor; chronic pain; tumour; neurological dysfunction;
XX basal ganglia; cholinergic interneuron; Parkinson's disease.
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX WO200131020-A1.
XX
XX 03-MAY-2001.
XX
XX 20-OCT-2000; 2000WO-US29064.
XX
XX 22-OCT-1999; 99US-0161159.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PT
```

```
XX Fitzgerald DJ, Iadarola MJ;
XX WPI; 2001-417560/44.
XX
XX Making cell toxin to treat chronic pain, by forming substance
XX P-pseudomonas exotoxin disulfide-linked conjugate, by reacting modified
XX exotoxin and substance P having additional cysteine residue at its
XX N-terminus .
XX
XX Example 1; Page 10; 54pp; English.
XX
XX The present sequence represents a modified substance P. The peptide is
XX used to produce a cell toxin. The cell toxin comprises a substance
XX P-pseudomonas exotoxin disulfide-linked conjugate. The cell toxin is
XX useful for ablating NK-1 receptor expressing cells, such as dorsal horn
XX cell, a stratum cell or a brain parenchyma cell, for treating chronic
XX pain in epineurium cells, perineurium cells, nerve ganglia, nerve
XX sheaths, nerve linings, meninges, pia mater cells, arachnoid membrane
XX cells, duramembrane cells, cells lining a joint or brain or spinal cord
XX parenchymal cells, without significantly affecting basal nociceptive
XX responses. The cell toxin is thus useful for treating chronic pain or
XX tumours that binds substance P. It is also useful for neurological
XX dysfunctions of the basal ganglia by targeting cholinergic interneurons
XX that express substance P e.g. Parkinson's disease.
XX
XX Sequence 12 AA;

Query Match 100.0%; Score 61; DB 22; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00032;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQQFFGLM 11
Db 2 rpqpqqffglm 12

RESULT 51
AAB98867
ID AAB98867 standard; Peptide; 12 AA.
XX
XX AAB98867;
XX
XX 14-AUG-2001 (first entry)
XX
XX Chimeric analgesic peptide #23.
XX
XX Opioid receptor binding; nociceptive receptor binding; analgesic;
XX pain; chimeric peptide.
XX
XX Synthetic.
XX
XX Key Location/Qualifiers
XX Modified-site 12
XX /label= OTHER
XX /note= "C-terminal amide"
XX
XX WO200130371-A2.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-US29789.
XX
XX 28-OCT-1999; 99US-0428692.
XX
XX (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.
XX
XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;
XX WPI; 2001-397593/42.
XX
XX New chimeric peptides used for treating pain comprise opiod receptor
XX
```

PT binding group and nociceptive receptor binding group -  
PS Claim 10; Page 15; 34pp; English.  
XX

CC The present invention describes a number of chimeric peptides comprising  
CC an opioid receptor binding moiety and a nociceptive receptor binding  
CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
CC opioid receptor based peptides alone, tolerance does not result from  
CC their long-term use. The present sequence is one of the peptides of the  
CC invention.  
XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 61; DB 22; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPOQFFGLM 11  
| | | | | | | | | |  
DB 1 rpkpqgffglm 11

RESULT 52  
AAB98870  
ID AAB98870 standard; Peptide; 12 AA.  
XX

AC AAB98870;

XX 14-AUG-2001 (first entry)

XX Chimeric analgesic peptide #26.

XX Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.  
XX

OS Synthetic.

FH Key Location/Qualifiers  
FT Modified-site 12  
FT /label= OTHER  
FT /note= "modified by OMe"

XX WO200130371-A2.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-US29789.

XX 28-OCT-1999; 99US-0428692.

XX (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;

XX WPI; 2001-397593/42.

XX New chimeric peptides used for treating pain comprise opioid receptor  
PT binding group and nociceptive receptor binding group -  
XX

PS Claim 10; Page 15; 34pp; English.

XX The present invention describes a number of chimeric peptides comprising  
CC an opioid receptor binding moiety and a nociceptive receptor binding  
CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
CC opioid receptor based peptides alone, tolerance does not result from  
CC their long-term use. The present sequence is one of the peptides of the  
CC invention.  
XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 61; DB 22; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPOQFFGLM 11  
| | | | | | | | | |  
DB 1 rpkpqgffglm 11

RESULT 53  
AAB98873  
ID AAB98873 standard; Peptide; 12 AA.  
XX

AC AAB98873;

XX 14-AUG-2001 (first entry)

XX Chimeric analgesic peptide #29.

XX Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.  
XX

OS Synthetic.

FH Key Location/Qualifiers  
FT Modified-site 12  
FT /label= OTHER  
FT /note= "modified by Oeth"

XX WO200130371-A2.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-US29789.

XX 28-OCT-1999; 99US-0428692.

XX (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;

XX WPI; 2001-397593/42.

XX New chimeric peptides used for treating pain comprise opioid receptor  
PT binding group and nociceptive receptor binding group -  
XX

PS Claim 10; Page 15; 34pp; English.

XX The present invention describes a number of chimeric peptides comprising  
CC an opioid receptor binding moiety and a nociceptive receptor binding  
CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
CC opioid receptor based peptides alone, tolerance does not result from  
CC their long-term use. The present sequence is one of the peptides of the  
CC invention.  
XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 61; DB 22; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPOQFFGLM 11  
| | | | | | | | | |  
DB 1 rpkpqgffglm 11

RESULT 54  
AAY03158  
ID AAY03158 standard; peptide; 13 AA.  
XX

AC AAY03158;

XX 10-JUN-1999 (first entry)

XX Substance P-Glycine-Lysine.  
DE Opioid peptide; opioid analgesia; enhancer; opioid anaesthesia;  
KW substance P.  
KW Synthetic.  
OS US5891842-A.  
XX  
PN  
XX  
PD 06-APR-1999.  
XX  
XX 12-APR-1996; 96US-0631434.  
XX  
XX 09-APR-1993; 93US-0044954.  
PR 12-APR-1996; 96US-0631434.  
XX  
XX (TUFT ) TUFTS COLLEGE.  
PA  
XX  
XX Cream RM;  
PI  
XX  
XX WPI; 1999-253906/21.  
DR  
XX  
XX Synergistic method for enhancing opioid analgesia and anaesthesia  
PT within a human  
PT  
XX  
XX Disclosure; Column 14; 20pp; English.  
PS  
XX  
XX This sequence represents substance P used in the method of the  
CC invention. The method is for enhancing opioid analgesia within a human  
CC subject for a duration of 15 minutes comprises concurrent administration  
CC of substance P, or one of its precursors. The method is used to elicit  
CC opioid analgesia and anaesthesia, either prior to or after the occurrence  
CC of a nociceptive event. The components have a synergistic effect. The  
CC method allows use of low doses of opioid that produce little or no  
CC physiological effect reducing conventional risks of toxicity and  
CC addiction, and allows the use of low doses of substance P and its related  
CC analogs that limit their in vivo physiological consequences. The  
CC analgesia is naloxone reversible allowing diminishment or complete  
CC elimination of opioid analgesia if desired and on demand. The treatment  
CC provides a durable analgesic effect, but only minimally disturbs and  
CC interrupts the normal metabolic processes of the body.  
XX  
XX  
SQ Sequence 13 AA;  
  
Query Match 100.0%; Score 61; DB 20; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00034;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 RPKPQQFFGLM 11  
Db 1 rpkpqffglm 11  
|||||  
  
RESULT 55  
AAG62770  
ID AAG62770 standard; peptide; 13 AA.  
XX  
XX  
AC AAG62770;  
XX  
XX  
DT 17-SEP-2001 (first entry)  
XX  
XX Amino acid sequence of substance P precursor.  
DE Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
KW Unidentified.  
XX  
XX WO200153336-A1.  
XX  
XX 26-JUL-2001.  
PD  
XX

PF 17-JAN-2001; 2001WO-US01529.  
XX  
XX  
PR 19-JAN-2000; 2000US-0489667.  
XX  
XX (ALLR ) ALLERGAN SALES INC.  
PA  
XX  
XX Donovan S;  
PI  
XX  
XX WPI; 2001-451900/48.  
DR  
XX  
XX Agent useful for treating pain comprises a clostridial neurotoxin (or  
PT component) attached to a targeting moiety -  
PT  
XX  
XX Disclosure; Page 62; 77pp; English.  
PS  
XX  
XX The specification describes an agent, comprising a clostridial neurotoxin  
CC attached to a targeting moiety, where the targeting moiety is selected  
CC from transmission compounds, and compounds substantially similar to the  
CC transmission compounds. The agent may be used for treating pain, where  
CC the clostridial neurotoxin component is derived from botulinum toxin  
CC selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
CC The targeting moiety comprises a light chain and an amine end segment of  
CC a heavy chain and comprises Substance P as the targeting moiety. The pain  
CC alleviating effects persist for 2-6 months. The present sequence  
CC represents substance P precursor, and is used in the course of the  
CC invention.  
XX  
XX  
SQ Sequence 13 AA;  
  
Query Match 100.0%; Score 61; DB 22; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.00034;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 RPKPQQFFGLM 11  
Db 1 rpkpqffglm 11  
|||||  
  
RESULT 56  
AAG62773  
ID AAG62773 standard; peptide; 13 AA.  
XX  
XX  
AC AAG62773;  
XX  
XX  
DT 17-SEP-2001 (first entry)  
XX  
XX Amino acid sequence of carboxy-ester substance P precursor.  
DE Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
KW Synthetic.  
XX  
XX  
XX Key Location/Qualifiers  
FH Modified-site 13 /note= "methylated residue"  
FT  
XX  
XX WO200153336-A1.  
PN  
XX  
XX 26-JUL-2001.  
PD  
XX  
XX  
PF 17-JAN-2001; 2001WO-US01529.  
XX  
XX  
PR 19-JAN-2000; 2000US-0489667.  
XX  
XX (ALLR ) ALLERGAN SALES INC.  
PA  
XX  
XX Donovan S;  
PI  
XX  
XX WPI; 2001-451900/48.  
DR  
XX  
XX Agent useful for treating pain comprises a clostridial neurotoxin (or  
PT component) attached to a targeting moiety -  
PT

XX Disclosure; Page 65; 77pp; English.

XX The specification describes an agent, comprising a clostridial neurotoxin

CC attached to a targeting moiety, where the targeting moiety is selected

CC from transmission compounds, and compounds substantially similar to the

CC transmission compounds. The agent may be used for treating pain, where

CC the clostridial neurotoxin component is derived from botulinum toxin

CC selected from botulinum types A, B, C, D, E, F, G and mixtures of these.

CC The targeting moiety comprises a light chain and an amine end segment of

CC a heavy chain and comprises Substance P as the targeting moiety. The pain

CC alleviating effects persist for 2-6 months. The present sequence

CC represents a substance P precursor, and is used in the course of the

CC invention.

XX SQ Sequence 13 AA;

Query Match , 100.0%; Score 61; DB 22; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00034;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11

Db 1 rpkpqffglm 11

|||||

RESULT 57

AAG62776

ID AAG62776 standard; peptide; 13 AA.

XX AC AAG62776;

XX DT 17-SEP-2001 (first entry)

XX DE Amino acid sequence of carboxy-ester substance P precursor.

XX KW Clostridial neurotoxin; pain; botulinum toxin; Substance P.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 13

FT /note= "ethylated residue"

XX PN WO200153336-A1.

XX PD 26-JUL-2001.

XX PF 17-JAN-2001; 2001WO-US01529.

XX PR 19-JAN-2000; 2000US-0489667.

XX PA (ALLR ) ALLERGAN SALES INC.

XX PI Donovan S;

XX WPI; 2001-451900/48.

XX Agent useful for treating pain comprises a clostridial neurotoxin (or

PT component) attached to a targeting moiety -

XX Disclosure; Page 68; 77pp; English.

XX The specification describes an agent, comprising a clostridial neurotoxin

CC attached to a targeting moiety, where the targeting moiety is selected

CC from transmission compounds, and compounds substantially similar to the

CC transmission compounds. The agent may be used for treating pain, where

CC the clostridial neurotoxin component is derived from botulinum toxin

CC selected from botulinum types A, B, C, D, E, F, G and mixtures of these.

CC The targeting moiety comprises a light chain and an amine end segment of

CC a heavy chain and comprises Substance P as the targeting moiety. The pain

CC alleviating effects persist for 2-6 months. The present sequence

CC represents a substance P precursor, and is used in the course of the

CC invention.

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 61; DB 22; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00034;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11

Db 1 rpkpqffglm 11

|||||

RESULT 58

AAB98868

ID AAB98868 standard; Peptide; 13 AA.

XX AC AAB98868;

XX DT 14-AUG-2001 (first entry)

XX DE Chimeric analgesic peptide #24.

XX KW Opioid receptor binding; nociceptive receptor binding; analgesic;

XX KW pain; chimeric peptide.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 13

FT /label= OTHER

FT /note= "C-terminal amide"

XX PN WO200130371-A2.

XX PD 03-MAY-2001.

XX PF 27-OCT-2000; 2000WO-US29789.

XX PR 28-OCT-1999; 99US-0428692.

XX PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

XX PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;

XX WPI; 2001-397593/42.

XX New chimeric peptides used for treating pain comprise opioid receptor

PT binding group and nociceptive receptor binding group -

XX Claim 10; Page 15; 34pp; English.

XX The present invention describes a number of chimeric peptides comprising

CC an opioid receptor binding moiety and a nociceptive receptor binding

CC moiety. These can be used as analgesics for the treatment of pain. Unlike

CC opioid receptor based peptides alone, tolerance does not result from

CC their long-term use. The present sequence is one of the peptides of the

CC invention.

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 61; DB 22; Length 13;

Best Local Similarity 100.0%; Pred. No. 0.00034;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11

Db 1 rpkpqffglm 11

|||||

```
RESULT 59
AAB98871
ID AAB98871 standard; Peptide; 13 AA.
XX
AC AAB98871;
XX
AC 14-AUG-2001 (first entry)
XX
DE Chimeric analgesic peptide #27.
XX
KW Opioid receptor binding; nociceptive receptor binding; analgesic;
KW pain; chimeric peptide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 13
FT /label= OTHER
FT /note= "modified by Ome"
XX
XX WO200130371-A2.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-US29789.
XX
XX 28-OCT-1999; 99US-0428692.
XX
XX (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.
XX
XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;
XX WPI; 2001-397593/42.
XX
XX New chimeric peptides used for treating pain comprise opioid receptor
binding group and nociceptive receptor binding group
XX
XX Claim 10; Page 15; 34pp; English.
XX
XX The present invention describes a number of chimeric peptides comprising
an opioid receptor binding moiety and a nociceptive receptor binding
moiety. These can be used as analgesics for the treatment of pain. Unlike
opioid receptor based peptides alone, tolerance does not result from
their long-term use. The present sequence is one of the peptides of the
invention.
XX
XX Sequence 13 AA;

Query Match 100.0%; Score 61; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11
DB 1 rpkpqffglm 11

RESULT 60
AAB98874
ID AAB98874 standard; Peptide; 13 AA.
XX
AC AAB98874;
XX
AC 14-AUG-2001 (first entry)
XX
DE Chimeric analgesic peptide #30.
XX
KW Opioid receptor binding; nociceptive receptor binding; analgesic;
KW pain; chimeric peptide.
XX
OS Synthetic.
XX

FH Key Location/Qualifiers
FT Modified-site 13
FT /label= OTHER
FT /note= "modified by Ome"
XX
XX WO200130371-A2.
XX
XX 03-MAY-2001.
XX
XX 27-OCT-2000; 2000WO-US29789.
XX
XX 28-OCT-1999; 99US-0428692.
XX
XX (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.
XX
XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;
XX WPI; 2001-397593/42.
XX
XX New chimeric peptides used for treating pain comprise opioid receptor
binding group and nociceptive receptor binding group
XX
XX Claim 10; Page 15; 34pp; English.
XX
XX The present invention describes a number of chimeric peptides comprising
an opioid receptor binding moiety and a nociceptive receptor binding
moiety. These can be used as analgesics for the treatment of pain. Unlike
opioid receptor based peptides alone, tolerance does not result from
their long-term use. The present sequence is one of the peptides of the
invention.
XX
XX Sequence 13 AA;

Query Match 100.0%; Score 61; DB 22; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11
DB 1 rpkpqffglm 11

RESULT 61
AAY03159
ID AAY03159 standard; peptide; 14 AA.
XX
AC AAY03159;
XX
XX 10-JUN-1999 (first entry)
XX
DE Substance P-Glycine-Lysine-Arginine.
XX
KW Opioid peptide; opioid analgesia; enhancer; opioid anaesthesia;
KW substance P.
XX
OS Synthetic.
XX
XX US5891842-A.
XX
XX 06-APR-1999.
XX
XX 12-APR-1996; 96US-0631434.
XX
XX 09-APR-1993; 93US-0044954.
XX
XX 12-APR-1996; 96US-0631434.
XX
XX (TUFT ) TUFTS COLLEGE.
XX
XX Kream RM;
XX
XX WPI; 1999-253906/21.
XX
```



PT Synergistic method for enhancing opioid analgesia and anaesthesia  
PT within a human  
XX  
XX Disclosure; Column 14; 20pp; English.  
XX  
CC This sequence represents substance P used in the method of the  
CC invention. The method is for enhancing opioid analgesia within a human  
CC subject for a duration of 15 minutes comprises concurrent administration  
CC of substance P, or one of its precursors. The method is used to elicit  
CC of opioid analgesia and anaesthesia, either prior to or after the occurrence  
CC of a nociceptive event. The components have a synergistic effect. The  
CC method allows use of low doses of opioid that produce little or no  
CC physiological effect reducing conventional risks of toxicity and  
CC addiction, and allows the use of low doses of substance P and its related  
CC analogs that limit their in vivo physiological consequences. The  
CC analgesia is naloxone reversible allowing diminishment or complete  
CC elimination of opioid analgesia if desired and on demand. The treatment  
CC provides a durable analgesic effect, but only minimally disturbs and  
CC interrupts the normal metabolic processes of the body.  
XX  
XX Sequence 14 AA;  
SQ

Query Match 100.0%; Score 61; DB 20; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11  
|||||

RESULT 62  
AAG62771  
ID AAG62771 standard; peptide; 14 AA.  
XX  
XX AAG62771;  
XX  
XX 17-SEP-2001 (first entry)  
XX  
XX Amino acid sequence of substance P precursor.  
XX Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
XX Unidentified.  
XX  
XX WO200153336-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01529.  
XX  
XX 19-JAN-2000; 2000US-0489667.  
XX (ALLR ) ALLERGAN SALES INC.  
XX Donovan S;  
XX WPI; 2001-451900/48.  
XX  
XX 26-JUL-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01529.  
XX  
XX 19-JAN-2000; 2000US-0489667.  
XX (ALLR ) ALLERGAN SALES INC.  
XX Donovan S;  
XX WPI; 2001-451900/48.  
XX  
XX Agent useful for treating pain comprises a clostridial neurotoxin (or  
XX component) attached to a targeting moiety -  
XX  
XX Disclosure; Page 63; 77pp; English.  
XX  
XX The specification describes an agent, comprising a clostridial neurotoxin  
XX attached to a targeting moiety, where the targeting moiety is selected  
XX from transmission compounds, and compounds substantially similar to the  
XX transmission compounds. The agent may be used for treating pain, where  
XX the clostridial neurotoxin component is derived from botulinum toxin  
XX selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
XX The targeting moiety comprises a light chain and an amine end segment of  
XX a heavy chain and comprises Substance P as the targeting moiety. The pain  
XX alleviating effects persist for 2-6 months. The present sequence of the  
XX represents a substance P precursor, and is used in the course of the  
XX invention.  
XX  
XX Sequence 14 AA;  
SQ

Query Match 100.0%; Score 61; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11  
|||||

RESULT 63  
AAG62774  
ID AAG62774 standard; peptide; 14 AA.  
XX  
XX AAG62774;  
XX  
XX 17-SEP-2001 (first entry)  
XX  
XX Amino acid sequence of carboxy-ester substance P precursor.  
XX Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
XX Synthetic.  
XX  
XX OS Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Modified-site 14 /note- "methylated residue"  
XX  
XX WO200153336-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01529.  
XX  
XX 19-JAN-2000; 2000US-0489667.  
XX (ALLR ) ALLERGAN SALES INC.  
XX Donovan S;  
XX WPI; 2001-451900/48.  
XX  
XX Agent useful for treating pain comprises a clostridial neurotoxin (or  
XX component) attached to a targeting moiety -  
XX  
XX Disclosure; Page 66; 77pp; English.  
XX  
XX The specification describes an agent, comprising a clostridial neurotoxin  
XX attached to a targeting moiety, where the targeting moiety is selected  
XX from transmission compounds, and compounds substantially similar to the  
XX transmission compounds. The agent may be used for treating pain, where  
XX the clostridial neurotoxin component is derived from botulinum toxin  
XX selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
XX The targeting moiety comprises a light chain and an amine end segment of  
XX a heavy chain and comprises Substance P as the targeting moiety. The pain  
XX alleviating effects persist for 2-6 months. The present sequence of the  
XX represents a substance P precursor, and is used in the course of the  
XX invention.  
XX  
XX Sequence 14 AA;  
SQ

Query Match 100.0%; Score 61; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11  
|||||

CC alleviating effects persist for 2-6 months. The present sequence  
CC represents substance P precursor, and is used in the course of the  
CC invention.  
XX  
XX Sequence 14 AA;  
SQ

Query Match 100.0%; Score 61; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11  
|||||

RESULT 63  
AAG62774  
ID AAG62774 standard; peptide; 14 AA.  
XX  
XX AAG62774;  
XX  
XX 17-SEP-2001 (first entry)  
XX  
XX Amino acid sequence of carboxy-ester substance P precursor.  
XX Clostridial neurotoxin; pain; botulinum toxin; Substance P.  
XX Synthetic.  
XX  
XX Key Location/Qualifiers  
XX Modified-site 14 /note- "methylated residue"  
XX  
XX WO200153336-A1.  
XX  
XX 26-JUL-2001.  
XX  
XX 17-JAN-2001; 2001WO-US01529.  
XX  
XX 19-JAN-2000; 2000US-0489667.  
XX (ALLR ) ALLERGAN SALES INC.  
XX Donovan S;  
XX WPI; 2001-451900/48.  
XX  
XX Agent useful for treating pain comprises a clostridial neurotoxin (or  
XX component) attached to a targeting moiety -  
XX  
XX Disclosure; Page 66; 77pp; English.  
XX  
XX The specification describes an agent, comprising a clostridial neurotoxin  
XX attached to a targeting moiety, where the targeting moiety is selected  
XX from transmission compounds, and compounds substantially similar to the  
XX transmission compounds. The agent may be used for treating pain, where  
XX the clostridial neurotoxin component is derived from botulinum toxin  
XX selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
XX The targeting moiety comprises a light chain and an amine end segment of  
XX a heavy chain and comprises Substance P as the targeting moiety. The pain  
XX alleviating effects persist for 2-6 months. The present sequence of the  
XX represents a substance P precursor, and is used in the course of the  
XX invention.  
XX  
XX Sequence 14 AA;  
SQ

Query Match 100.0%; Score 61; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11  
|||||

Db 1 rpkpqffglm 11  
|||||

## RESULT 64

AAG62777  
ID AAG62777 standard; peptide; 14 AA.

XX AC AAG62777;

DT 17-SEP-2001 (first entry)

XX Amino acid sequence of carboxy-ester substance P precursor.

XX Clostridial neurotoxin; pain; botulinum toxin; Substance P.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 14

FT /note= "ethylated residue"

XX WO200153336-A1.

XX 26-JUL-2001.

XX 17-JAN-2001; 2001WO-US01529.

XX 19-JAN-2000; 2000US-0489667.

XX (ALLR ) ALLERGAN SALES INC.

XX Donovan S;

XX WPI; 2001-451900/48.

XX Agent useful for treating pain comprises a clostridial neurotoxin (or

XX component) attached to a targeting moiety

XX Disclosure; Page 69; 77pp; English.

XX The specification describes an agent, comprising a clostridial neurotoxin  
XX attached to a targeting moiety, where the targeting moiety is selected  
XX from transmission compounds, and compounds substantially similar to the  
XX transmission compounds. The agent may be used for treating pain, where  
XX the clostridial neurotoxin component is derived from botulinum toxin  
XX selected from botulinum types A, B, C, D, E, F, G and mixtures of these.  
XX The targeting moiety comprises a light chain and an amine end segment of  
XX a heavy chain and comprises Substance P as the targeting moiety. The pain  
XX alleviating effects persist for 2-6 months. The present sequence  
XX represents a substance P precursor, and is used in the course of the  
XX invention.

XX Sequence 14 AA;

## Query Match

Best Local Similarity 100.0%; Score 61; DB 22; Length 14;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11

Db 1 rpkpqffglm 11  
|||||

## RESULT 65

AAB98869  
ID AAB98869 standard; peptide; 14 AA.

XX AC AAB98869;

DT 14-AUG-2001 (first entry)

XX

DE Chimeric analgesic peptide #25.

XX Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 14

FT /label= OTHER

FT /note= "C-terminal amide"

XX WO200130371-A2.

XX 03-MAY-2001.

XX 27-OCT-2000; 2000WO-US29789.

XX 28-OCT-1999; 99US-0428692.

XX (NEW-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.

XX Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;

XX WPI; 2001-397593/42.

XX New chimeric peptides used for treating pain comprise opioid receptor  
XX binding group and nociceptive receptor binding group

XX Claim 10; Page 15; 34pp; English.

XX The present invention describes a number of chimeric peptides comprising  
XX an opioid receptor binding moiety and a nociceptive receptor binding  
XX moiety. These can be used as analgesics for the treatment of pain. Unlike  
XX opioid receptor based peptides alone, tolerance does not result from  
XX their long-term use. The present sequence is one of the peptides of the  
XX invention.

XX Sequence 14 AA;

## Query Match

Best Local Similarity 100.0%; Score 61; DB 22; Length 14;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11

Db 1 rpkpqffglm 11  
|||||

## RESULT 66

AAB98872

ID AAB98872 standard; Peptide; 14 AA.

XX AC AAB98872;

DT 14-AUG-2001 (first entry)

DE Chimeric analgesic peptide #28.

XX Opioid receptor binding; nociceptive receptor binding; analgesic;  
KW pain; chimeric peptide.

XX OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 14

FT /label= OTHER

FT /note= "modified by OMe"

XX WO200130371-A2.

XX 03-MAY-2001.

XX PF 27-OCT-2000; 2000WO-US29789.  
XX PR 28-OCT-1999; 99US-0428692.  
XX PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX WPI; 2001-397593/42.  
XX DR New chimeric peptides used for treating pain comprise opioid receptor  
XX PT binding group and nociceptive receptor binding group  
XX PS Claim 10; Page 15; 34pp; English.  
XX CC The present invention describes a number of chimeric peptides comprising  
XX CC an opioid receptor binding moiety and a nociceptive receptor binding  
XX CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
XX CC opioid receptor based peptides alone, tolerance does not result from  
XX CC their long-term use. The present sequence is one of the peptides of the  
XX CC invention.  
XX SQ Sequence 14 AA;  
DR WPI; 2001-397593/42.  
XX CC New chimeric peptides used for treating pain comprise opioid receptor  
XX PT binding group and nociceptive receptor binding group  
XX PS Claim 10; Page 15; 34pp; English.

XX PF 27-OCT-2000; 2000WO-US29789.  
XX PR 28-OCT-1999; 99US-0428692.  
XX PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX WPI; 2001-397593/42.  
XX DR New chimeric peptides used for treating pain comprise opioid receptor  
XX PT binding group and nociceptive receptor binding group  
XX PS Claim 10; Page 15; 34pp; English.

Query Match 100.0%; Score 61; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11  
RESULT 67  
AAB98875  
ID AAB98875 standard; Peptide; 14 AA.  
XX AC AAB98875;  
XX DT 14-AUG-2001 (first entry)  
XX DE Chimeric analgesic peptide #31.  
XX KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
XX KW pain; chimeric peptide.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Modified-site 14  
FT /label= OTHER  
FT /note= "modified by Oeth"  
XX PN WO200130371-A2.  
XX PD 03-MAY-2001.  
XX PF 27-OCT-2000; 2000WO-US29789.  
XX PR 28-OCT-1999; 99US-0428692.  
XX PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX WPI; 2001-397593/42.  
XX DR New chimeric peptides used for treating pain comprise opioid receptor  
XX PT binding group and nociceptive receptor binding group  
XX PS Claim 10; Page 15; 34pp; English.

Query Match 100.0%; Score 61; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11  
RESULT 67  
AAB98875  
ID AAB98875 standard; Peptide; 14 AA.  
XX AC AAB98875;  
XX DT 14-AUG-2001 (first entry)  
XX DE Chimeric analgesic peptide #31.  
XX KW Opioid receptor binding; nociceptive receptor binding; analgesic;  
XX KW pain; chimeric peptide.  
XX OS Synthetic.  
XX FH Key Location/Qualifiers  
FT Modified-site 14  
FT /label= OTHER  
FT /note= "modified by Oeth"  
XX PN WO200130371-A2.  
XX PD 03-MAY-2001.  
XX PF 27-OCT-2000; 2000WO-US29789.  
XX PR 28-OCT-1999; 99US-0428692.  
XX PA (NEWE-) NEW ENGLAND MEDICAL CENT HOSPITALS INC.  
XX PI Carr DB, Lipkowski AW, Kream R, Misicka-Kesik A;  
XX WPI; 2001-397593/42.  
XX DR New chimeric peptides used for treating pain comprise opioid receptor  
XX PT binding group and nociceptive receptor binding group  
XX PS Claim 10; Page 15; 34pp; English.

XX CC The present invention describes a number of chimeric peptides comprising  
XX CC an opioid receptor binding moiety and a nociceptive receptor binding  
XX CC moiety. These can be used as analgesics for the treatment of pain. Unlike  
XX CC opioid receptor based peptides alone, tolerance does not result from  
XX CC their long-term use. The present sequence is one of the peptides of the  
XX CC invention.  
XX SQ Sequence 14 AA;  
DR WPI; 2001-397593/42.  
XX CC New chimeric peptides used for treating pain comprise opioid receptor  
XX PT binding group and nociceptive receptor binding group  
XX PS Claim 10; Page 15; 34pp; English.

Query Match 100.0%; Score 61; DB 22; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.00037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX QY 1 RPKPQQFFGLM 11  
DB 1 rpkpqffglm 11  
RESULT 68  
AAB91440  
ID AAB91440 standard; Peptide; 14 AA.  
XX AC AAB91440;  
XX DT 22-JUN-2001 (first entry)  
XX DE Tachykinins peptide SEQ ID NO:616.  
XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
XX KW blood component; modification; succinimide; maleimido group; amino;  
XX KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX OS Homo sapiens.  
XX OS Synthetic.  
XX PN WO200069900-A2.  
XX PD 23-NOV-2000.  
XX PF 17-MAY-2000; 2000WO-US13576.  
XX PR 17-MAY-1999; 99US-0134406.  
XX PR 10-SEP-1999; 99US-0153406.  
XX PR 15-OCT-1999; 99US-0159783.  
XX PA (CONJ-) CONJUCHEM INC.  
XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;  
XX WPI; 2001-112059/12.  
XX PT Modifying and attaching therapeutic peptides to albumin prevents  
XX PT peptidase degradation, useful for increasing length of in vivo activity  
XX PS Disclosure; Page 400; 733pp; English.  
XX CC The present invention describes a modified therapeutic peptide (I)  
XX CC comprising a therapeutically active amino acid region (III) and a  
XX CC reactive group (II) (e.g. succinimide and maleimido groups) attached to  
XX CC a less therapeutically active amino acid region (IV), which covalently  
XX CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX CC factors and neurotransmitters, to protect them from peptidase activity  
XX CC in vivo for the treatment of various disorders. Endogenous therapeutic  
XX CC peptides are not suitable as drug candidates as they require frequent  
XX CC administration due to rapid degradation by peptidases in the body.  
XX CC Modifying and attaching therapeutic peptides to albumin prevents or  
XX CC reduces the action of peptidases to increase length of activity (half  
XX CC life) and specificity as bonding to large molecules decreases

CC Intracellular uptake and interference with physiological processes.  
 CC AAB90829 to AAB92441 represent peptides which can be used in the  
 CC exemplification of the present invention.

XX Sequence 14 AA;

Query Match 100.0%; Score 61; DB 22; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 0.00037;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
 |||||  
 Db 1 rpkpqgffglm 11

RESULT 69  
 AAB06258  
 ID AAB06258 standard; peptide; 20 AA.

XX AAB06258;

DT 16-OCT-2000 (first entry)

XX Substance P analogue #2.

XX Substances P; SP; neurokinin-1 receptor; NK-1R; nociception; NTE-SAP;  
 KW Saporin; SAP; analgesic; anti-inflammatory; neuroprotective;  
 KW anti-asthmatic; anti-allergic; dermatological; anti-ulcer;  
 KW tranquiliser; immunosuppressive; anti-migraine; cytostatic;  
 KW substance P antagonist; cytotoxic; ribosome inactivator;  
 KW prostaglandin antagonist; cancer; respiratory disease; asthma;  
 KW allergic rhinitis; ophthalmic disease; conjunctivitis;  
 KW allergic dermatitis; psoriasis; ulcerative colitis; Crohn's disease;  
 KW gastrointestinal disorder; anxiety; psychosis; rheumatoid arthritis;  
 KW carcinoma; lupus erythematosus conjunctivitis.

XX Synthetic.

Key Location/Qualifiers  
 Modified-site 20 /note= "C-terminal amide"

US6063758-A.

16-MAY-2000.

09-JUL-1997; 97US-0890157.

09-JUL-1997; 97US-0890157.

(ADTA-) ADVANCED TARGETING SYSTEMS INC.

Lappi DA, Wiley RG;

WPI; 2000-430049/37.

New conjugates comprising substance P or its analog, and a  
 ribosome-inactivating protein (for example saporin), for alleviating  
 pain and treating disorders associated with neurokinin-1 receptor -

Claim 1; Column 2; 21pp; English.

The present sequence is an analogue of substance P (SP). SP, which binds  
 to the neurokinin-1 receptor (NK-1R), is best known for its role in  
 nociception. It is secreted by small unmyelinated C-fibres of the  
 peripheral nervous system that are thought to be primary nociceptive  
 neurons. The present sequence may be conjugated to Saporin (SAP), a  
 ribosome-inactivating protein, to produce NTE-SAP. The conjugate may be  
 used to control chronic pain by specifically targeting cells having NK1  
 receptors, and inhibiting proliferation of or causing death of these  
 cells. It may also be used to treat NK-1R-associated disorders  
 including respiratory conditions (e.g. asthma, allergic rhinitis),

CC ophthalmic conditions (e.g. conjunctivitis), cutaneous conditions (e.g.  
 CC allergic dermatitis, psoriasis), intestinal conditions (e.g. ulcerative  
 CC colitis, Crohn's disease), gastrointestinal disorders, central nervous  
 CC system disorders (e.g. anxiety, psychosis), inflammatory diseases (e.g.  
 CC rheumatoid arthritis), proliferative conditions (e.g. carcinoma),  
 CC disorders related to immune enhancement or suppression (e.g. lupus  
 CC erythematosus conjunctivitis), and especially migraine.

SQ Sequence 20 AA;

Query Match 100.0%; Score 61; DB 21; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.00052;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
 |||||  
 Db 10 rpkpqgffglm 20

RESULT 70

AAP70431

ID AAP70431 standard; protein; 129 AA.

XX AAP70431;

XX 17-JAN-1991 (first entry)

XX Human beta-preprotachykinin.

XX Preprotachykinin; substance P; neurokinin A; tachykinin;

XX Homo sapiens.

Key Location/Qualifiers

Region 20..56

FT /label=claimed polypeptide

Region 1..126

FT /label=claimed polypeptide

Region 111..126

FT /label=claimed polypeptide

XX WO8707643-A.

XX 17-DEC-1987.

XX 03-JUN-1987; 87WO-GB00382.

XX 03-JUN-1986; 86GB-0013431.

XX (RESE ) RESEARCH CORPORATION LTD.

XX Harmar AJ, Pascall J, Mckeown A;

XX WPI; 1987-362730/51.

XX N-PSDB; AAN70688.

XX New DNA sequence coding for the new polypeptide preprotachykinin -  
 a precursor for substance P, etc., useful as neurotransmitters,  
 diagnostic reagents, etc.

XX Claim 1; page 15; 25pp; English.

XX Beta-preprotachykinin includes sequences identical to tachykinins, eg  
 CC substance P, neurokinin A, or other biologically active peptides, eg  
 CC neuropeptide K. These peptides are, eg neurotransmitters, hormones,  
 CC analgesics and anti-inflammatories. The polypeptides can be used  
 CC as reagents in RIA, eg to monitor or diagnose carcinoid syndrome.

XX Sequence 129 AA;

Query Match 100.0%; Score 61; DB 8; Length 129;

Best Local Similarity 100.0%; Pred. No. 0.0031;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOFFGLM 11  
|||||  
Db 58 rpkpqffgilm 68

## RESULT 71

AAG99353  
ID AAG99353 standard; Protein; 129 AA.

XX AC AAG99353;

XX 25-SEP-2001 (first entry)

XX Human atypical tachykinin protein fragment SEQ ID NO: 63.

XX Atypical tachykinin; ATT; human; hypertension.

XX Homo sapiens.

XX WO200146415-A1.

XX 28-JUN-2001.

XX 21-DEC-2000; 2000WO-JP09083.

XX 21-DEC-1999; 99JP-0362638.

XX 10-MAR-2000; 2000JP-0066714.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Itoh Y, Nishi K, Kitada C, Inatomi N;

XX WPI; 2001-441676/47.

XX Atypical tachykinin peptides of human origin and DNA encoding them for  
screening potential agents for treatment of hypertension

XX Example 14; Page 143; 153pp; Japanese.

XX The present invention relates to atypical tachykinin proteins of human  
origin and their esters, amides, salts and partial peptides. These can be  
used in the treatment, prevention and diagnosis of hypertension. The  
present sequence is a protein fragment described in the exemplification  
of the invention.

XX Sequence 129 AA;

Query Match 100.0%; Score 61; DB 22; Length 129;  
Best Local Similarity 100.0%; Pred. No. 0.0031;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOFFGLM 11  
|||||  
Db 58 rpkpqffgilm 68

## RESULT 72

AAW16339  
ID AAW16339 standard; Protein; 401 AA.

XX AC AAW16339;

XX 05-SEP-1997 (first entry)

XX DAB389-SP-Gly fusion toxin.

XX DAB389-SP-Gly; amidated polypeptide binding ligand; drug delivery;  
KW diphtheria toxin; substance P; cancer; therapy.

## OS Synthetic.

XX WO9713410-A1.

XX 17-APR-1997.

XX 11-OCT-1996; 96WO-US16237.

XX 13-OCT-1995; 95US-0005431.

XX (BOST-) BOSTON MEDICAL CENT CORP.

XX Fisher CE, Leeman SE, Murphy JR, Vanderspek JC;

XX WPI; 1997-235583/21.

XX N-PSDB; AAF63359.

XX Hybrid molecule for targeting compound, especially a toxin, into  
cells - includes polypeptide able to transport the compound across  
cytoplasmic membranes and amidated ligand, useful for treatment of  
cancer

XX Example 1; Page 22-23; 51pp; English.

XX DAB389-SP-Gly (AAW16339) is a hybrid toxin comprising DAB389 (i.e.  
amino acids 1-386 plus His-484 and Ala-485 of mature diphtheria  
toxin) fused to C-terminal glycine-extended substance P. It was  
expressed in E. coli HMS174 (DE3) transformants using a vector  
that carried DAB389-SP-Gly DNA (see also AAT63359). The fusion  
protein was then amidated using peptidylglycine-alpha-amidating  
monooxygenase. The amidated fusion protein used to target DAB389  
toxin to specific cells contg. substance P receptors, esp. cancer  
cells. For human IM9 (chronic myelogenous leukaemia) cells contg.  
approx. 4000 substance P receptors per cell, the IC50 for amidated  
DAB389-SP-Gly was 18 pM.

XX Sequence 401 AA;

Query Match 100.0%; Score 61; DB 18; Length 401;  
Best Local Similarity 100.0%; Pred. No. 0.0092;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOFFGLM 11  
|||||  
Db 390 rpkpqffgilm 400

## RESULT 73

AAW26510  
ID AAW26510 standard; Protein; 487 AA.

XX AC AAW26510;

XX 06-JAN-1998 (first entry)

XX Amyloid precursor protein substrate APP-REP 751.

XX Amyloid precursor protein; APP; beta-amyloid protein; BAP;  
KW substrate; mutein; secretase; Alzheimer's disease; human;  
APP-REP 751; pCLL621.

XX Chimeric Homo sapiens.

XX Chimeric synthetic.

XX Key Location/Qualifiers

XX 362..372

XX /label= SP

XX /note= "substance P reporter epitope"

XX 389..430

XX /label= BAP

XX /note= "beta-amyloid protein"

XX Cleavage-site 404..405

FT FT /note= "secretase cleavage site"  
 FT FT 417..440  
 FT FT /label= Transmembrane  
 XX  
 PN US5656477-A.  
 XX  
 PD 12-AUG-1997.  
 XX  
 XX 01-MAY-1992; 92US-0877675.  
 XX  
 XX 20-SEP-1993; 93US-0123659.  
 PR 01-MAY-1992; 92US-0877675.  
 XX  
 PA (AMCY ) AMERICAN CYANAMID CO.  
 XX  
 XX Jacobsen JS, Vitek MP;  
 PI  
 DR WPI; 1997-414594/38.  
 DR P-PSDB; AAT87083.  
 XX  
 PT Nucleic acid encoding amyloid precursor mutin(s) - comprising  
 PT reporter gene and coding sequence, for identifying compounds which  
 PT modify the activity of proteolytic enzymes which cleave APP  
 XX  
 PS Disclosure; Fig 8; 84pp; English.

CC This polypeptide, designated APP-REP 761, comprises an amyloid  
 CC precursor protein (APP) that has a 276-amino acid deletion of the  
 CC native APP and which carries a Substance P epitope markers placed  
 CC N-terminal to the beta-amyloid protein (BAP) domain. APP-REP 751  
 CC can be used in a claimed method for screening for a compound which  
 CC reduces the formation of beta-amyloid protein, determined by  
 CC measuring the amount of marker in a medium containing transfected  
 CC cells. The method is used to detect compounds which inhibit the  
 CC activity of proteolytic enzymes which cleave APP to generate BAP  
 CC fragments. Such compounds can be used in the treatment of e.g.  
 CC Alzheimer's disease. The deletion of a 276 amino acid portion of  
 CC APP distinguishes the construct from endogenously expressed APP,  
 CC and beneficially increases the resolution of APP-REP fragments  
 CC resulting from the proteolytic cleavage by secretase or other  
 CC amyloidogenic, BAP-generating cleavage events.

XX Sequence 487 AA;

Query Match 100.0%; Score 61; DB 18; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
 |||||  
 Db 362 rpkpqffglm 372

RESULT 74  
 AAW26394  
 ID AAW26394 standard; Protein; 487 AA.

XX AAW26394;

DT 15-DEC-1997 (first entry)

DE Amyloid precursor protein substrate APP-REP 751.

KW Amyloid precursor protein; APP; beta-amyloid protein; BAP;  
 KW substrate; mutin; secretase; Alzheimer's disease; human;  
 KW APP-REP 751; pCLL621.

XX Chimeric Homo sapiens;  
 OS Chimeric synthetic.

PH Key Location/Qualifiers  
 FT Peptide 362..372

FT FT /label= SP  
 FT FT /note= "substance P reporter epitope"  
 FT FT 389..430  
 FT FT /label= BAP  
 FT FT /note= "Beta-amyloid protein"  
 FT FT 404..405  
 FT FT /note= "secretase cleavage site"  
 FT FT 417..440  
 FT FT /label= Transmembrane  
 XX  
 PN US5652092-A.  
 XX  
 PD 29-JUL-1997.  
 XX  
 XX 01-MAY-1992; 92US-0877675.  
 XX  
 XX 20-SEP-1993; 93US-0123659.  
 PR 01-MAY-1992; 92US-0877675.  
 PR 05-JUN-1995; 95US-0462859.  
 XX  
 PA (AMCY ) AMERICAN CYANAMID CO.  
 XX  
 XX Jacobsen JS, Vitek MP;  
 PI  
 DR WPI; 1997-392937/36.  
 DR N-PSDB; AAT84562.  
 XX  
 PT Screening for compounds which reduce beta-amyloid protein formation  
 PT - using cells which express a construct encoding a marker and an  
 PT amyloid precursor mutin derived from APP isoforms  
 XX  
 PS Disclosure; Fig 8; 84pp; English.

CC This polypeptide, designated APP-REP 761, comprises an amyloid  
 CC precursor protein (APP) that has a 276-amino acid deletion of the  
 CC native APP and which carries a Substance P epitope markers placed  
 CC N-terminal to the beta-amyloid protein (BAP) domain. APP-REP 751  
 CC can be used in a claimed method for screening for a compound which  
 CC reduces the formation of beta-amyloid protein, determined by  
 CC measuring the amount of marker in a medium containing transfected  
 CC cells. The method is used to detect compounds which inhibit the  
 CC activity of proteolytic enzymes which cleave APP to generate BAP  
 CC fragments. Such compounds can be used in the treatment of e.g.  
 CC Alzheimer's disease. The deletion of a 276 amino acid portion of  
 CC APP distinguishes the construct from endogenously expressed APP,  
 CC and beneficially increases the resolution of APP-REP fragments  
 CC resulting from the proteolytic cleavage by secretase or other  
 CC amyloidogenic, BAP-generating cleavage events.

XX Sequence 487 AA;

Query Match 100.0%; Score 61; DB 18; Length 487;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
 |||||  
 Db 362 rpkpqffglm 372

RESULT 75  
 AAW44745  
 ID AAW44745 standard; Protein; 487 AA.

XX AAW44745;

XX 01-JUN-1998 (first entry)

DE APP-REP 751 protein from pCLL621.

KW Amyloid precursor protein; APP; APP 751 isoform; deletion; substrate P;  
 KW epitope; Met-enkephalin; detection; secretase; beta-amyloid protein; BAP;

KW Alzheimer's disease; cleavage.

XX Homo sapiens.  
OS Synthetic.

PN US5693478-A.

PD 02-DEC-1997.

XX 05-JUN-1995; 95US-0464247.

XX 20-SEP-1993; 93US-0123659.

PR 01-MAY-1992; 92US-0877675.

XX 05-JUN-1995; 95US-0464247.

PA (AMCY ) AMERICAN CYANAMID CO.

PI Jacobsen JS, Vitek MP;

XX N-PSDB; AAV05850.

PT Amyloid precursor muterin reporter molecule assay containing antibody  
PT recognised marker - used to study pathways associated with

PT Alzheimer's disease

XX Disclosure; Fig 8; 84pp; English.

XX This is the amino acid sequence of a novel amyloid precursor protein  
CC (APP) designated APP-REP 751, contained in construct pCLL621. The  
CC sequence comprises a mutant version of the APP 751 isoform of human APP  
CC which contains a deletion of 276 amino acids from the central region.  
CC The deleted region is replaced by a substrate P reporter epitope  
CC sequence (RPKPQQFFGLM). In contrast to the APP-REP 751 encoded by the  
CC construct pCLL602 (AAW44744), this sequence does not contain a  
CC Met-enkephalin reporter epitope (YGGFM) fused at the C-terminus of the  
CC coding sequence. The shorter protein is generated for ease of detection  
CC based on size difference with the wild type APP protein and also by  
CC detection of the reporter epitopes. The mutant protein can be used in a  
CC method to study secretase and beta-amyloid protein (BAP)-generating  
CC pathways associated with Alzheimer's disease by studying proteolytic  
CC cleavage of the reporter polypeptides.

XX Sequence 487 AA;

Query Match 100.0%; Score 61; DB 19; Length 487;

Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11

Db ||||| |||||

362 rpkpqqffglm 372

RESULT 76

AAW42979

ID AAW42979 standard; Protein; 487 AA.

XX AAW42979;

XX 01-MAY-1998 (first entry)

DE Amyloid precursor protein mutant APP-ARP 751.

XX Beta-amyloid peptide; BAP; extracellular BAP plaque;  
KW cerebrovascular deposit; Alzheimers disease; Downs syndrome;  
KW amyloid precursor protein; APP; secretase; BAP aggregation;  
KW abnormal proteolytic cleavage.

XX Synthetic.

OS Homo sapiens.

XX

PN US5703209-A.

XX 30-DEC-1997.

XX 05-JUN-1995; 95US-0464248.

XX 20-SEP-1993; 93US-0123659.

PR 01-MAY-1992; 92US-0877675.

XX (AMCY ) AMERICAN CYANAMID CO.

XX Jacobsen JS, Vitek MP;

XX WPI; 1998-076482/07.

DR N-PSDB; AAV04866.

XX Amyloid precursor protein fusion polypeptides - comprising APP

PT fragment and marker, useful for research and drug screening

XX Disclosure; Fig 8A-Q; 84pp; English.

XX The present sequence represents an amyloid precursor protein (APP),  
CC which has a deletion of 276 amino acids to within 15 amino acids of the  
CC beta-amyloid peptide (BAP) domain. The protein also contains the Abnormal  
CC accumulation of extracellular BAP in plaques and cerebrovascular deposits  
CC is characteristic in brains of individuals suffering from Alzheimers  
CC disease and Downs syndrome. BAP is a poorly soluble, self-aggregating  
CC protein which is derived from a larger amyloid precursor protein (APP).  
CC APP is expressed as an integral membrane protein, and is cleaved by  
CC secretase, between BAP 16lys and 17leu. Cleavage at this site precludes  
CC amyloidogenesis and results in the release of the amino-terminal APP  
CC fragment. Three major isoforms of APP exist: APP-695, APP-751 and  
CC APP-770. These isoforms are derived by alternative splicing. APP-RRP 751  
CC is constructed by ligating restriction fragments representing N- and  
CC C-terminal APP-751 cDNA and substrate P reporter epitope sequences.  
CC APP can be used as a substrate for studying abnormal proteolytic cleavage  
CC which results in the release of BAP, and also to screen for drugs that  
CC will inhibit such cleavage.

XX Sequence 487 AA;

Query Match 100.0%; Score 61; DB 19; Length 487;

Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11

Db ||||| |||||

362 rpkpqqffglm 372

RESULT 77

AAW45229

ID AAR45229 standard; Protein; 492 AA.

XX AAR45229;

XX 20-JUN-1994 (first entry)

DE APP-REP 751 amyloid precursor protein/reporter protein.

XX Amyloid precursor protein; APP; beta amyloid protein; BAP;  
KW detection; Alzheimer's disease; Down's syndrome.

XX Homo sapiens.

XX AU9338358-A.

PN 04-NOV-1993.

XX 03-MAY-1993; 93AU-0038358.

XX 01-MAY-1992; 92US-0877675.

PR





FT Domain 417..440  
 FT /label= Transmembrane  
 FT Peptide 488..492  
 FT /label= ME  
 FT /note= "Met-enkephalin reporter epitope"  
 XX  
 PN US5652092-A.  
 XX  
 XX 29-JUL-1997.  
 PD  
 XX 01-MAY-1992; 92US-0877675.  
 PF  
 XX 20-SEP-1993; 93US-0123659.  
 PR 01-MAY-1992; 92US-0877675.  
 PR 05-JUN-1995; 95US-0462859.  
 XX  
 PA (AMCY ) AMERICAN CYANAMID CO.  
 XX  
 PI Jacobsen JS, Vitek MP;  
 XX WPI: 1997-392937/36.  
 DR N-PSDB; AAT84561.  
 DR  
 XX Screening for compounds which reduce beta-amyloid protein formation  
 PT - using cells which express a construct encoding a marker and an  
 PT amyloid precursor mutin derived from APP isoforms  
 XX  
 PS Disclosure; Fig 7; 84pp; English.  
 XX  
 CC This polypeptide, designated APP-REP 761, comprises an amyloid  
 CC precursor protein (APP) that has a 276-amino acid deletion of the  
 CC native APP and which carries Substance P and Met-enkephalin epitope  
 CC markers placed, respectively, on the N-terminal and C-terminal  
 CC sites of the beta-amyloid protein (BAP) domain. APP-REP 751 can  
 CC be used in a claimed method for screening for a compound which  
 CC reduces the formation of beta-amyloid protein, determined by  
 CC measuring the amount of marker in a medium containing transfected  
 CC cells. The method is used to detect compounds which inhibit the  
 CC activity of proteolytic enzymes which cleave APP to generate BAP  
 CC fragments. Such compounds can be used in the treatment of e.g.  
 CC Alzheimer's disease. The deletion of a 276 amino acid portion of  
 CC APP distinguishes the construct from endogenously expressed APP,  
 CC and beneficially increases the resolution of APP-REP fragments  
 CC resulting from the proteolytic cleavage by secretase or other  
 CC amyloidogenic, BAP-generating cleavage events.  
 XX  
 SQ Sequence 492 AA;

Query Match 100.0%; Score 61; DB 18; Length 492;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
 |||||  
 Db 362 rpkpgqffgglm 372

RESULT 80  
 AAW44744  
 ID AAW44744 standard; Protein; 492 AA.  
 XX  
 AC AAW44744;  
 XX  
 DT 01-JUN-1998 (first entry)  
 XX  
 DE APP-REP 751 protein from pCLL602.  
 XX  
 KW Amyloid precursor protein; APP; APP 751 isoform; deletion; substrate P;  
 KW epitope; Met-enkephalin; detection; secretase; beta-amyloid protein; BAP;  
 KW Alzheimer's disease; cleavage.  
 XX  
 OS Homo sapiens.

OS Synthetic.  
 XX US5693478-A.  
 PN  
 XX 02-DEC-1997.  
 PD  
 XX 05-JUN-1995; 95US-0464247.  
 PF  
 XX 20-SEP-1993; 93US-0123659.  
 PR 01-MAY-1992; 92US-0877675.  
 PR 05-JUN-1995; 95US-0464247.  
 XX  
 PA (AMCY ) AMERICAN CYANAMID CO.  
 XX  
 PI Jacobsen JS, Vitek MP;  
 XX WPI: 1998-031744/03.  
 DR N-PSDB; AAV05849.  
 DR  
 XX Amyloid precursor mutin reporter molecule assay containing antibody  
 PT recognised marker - used to study pathways associated with  
 PT Alzheimer's disease  
 PT  
 XX Disclosure; Fig 7; 84pp; English.  
 PS  
 XX This is the amino acid sequence of a novel amyloid precursor protein  
 CC (APP) designated APP-REP 751, contained in construct pCLL602. The  
 CC sequence comprises a mutant version of the APP 751 isoform of human APP  
 CC which contains a deletion of 276 amino acids from the central region.  
 CC The deleted region is replaced by a substrate P reporter epitope sequence  
 CC (RPKPPQFFGLM) and a Met-enkephalin reporter epitope (YGGFM) is fused at  
 CC the C-terminus. The shorter protein is generated for ease of detection  
 CC based on size difference with the wild type APP protein and also by  
 CC detection of the reporter epitopes. The mutant protein can be used in  
 CC a method to study secretase and beta-amyloid protein (BAP)-generating  
 CC pathways associated with Alzheimer's disease by studying proteolytic  
 CC cleavage of the reporter polypeptides.  
 XX  
 SQ Sequence 492 AA;

Query Match 100.0%; Score 61; DB 19; Length 492;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
 |||||  
 Db 362 rpkpgqffgglm 372

RESULT 81  
 AAW42978  
 ID AAW42978 standard; Protein; 492 AA.  
 XX  
 AC AAW42978;  
 XX  
 DT 01-MAY-1998 (first entry)  
 XX  
 DE Amyloid precursor protein mutant APP-APP 751.  
 XX  
 KW Beta-amyloid peptide; BAP; extracellular BAP plaque;  
 KW cerebrovascular deposit; Alzheimers disease; Downs syndrome;  
 KW amyloid precursor protein; APP; secretase; BAP aggregation;  
 KW abnormal proteolytic cleavage.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1..487  
 FT /note= "APP-APP 751"  
 FT Peptide 488..492  
 FT /note= "Met-enkephalin reporter epitope"



PT Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX  
PS Claim 10; Page 21; 35pp; English.  
XX  
CC The peptide is the tachykinin agonist substance P with a Norleucine  
CC residue substituted at position 11. The peptide was synthesised  
CC by standard solid phase synthesis. An N-terminal deleted peptide  
CC (7-11) with the same substitution was also synthesised. Neuronal  
CC accumulation of beta-amyloid may be treated by administration of  
CC tachykinin agonists. The peptides can reduce the neurotoxic effects  
CC of a beta-amyloid related polypeptide on cultured neurons. The  
CC peptide and its analogues are useful for controlling diseases  
CC characterised by beta amyloid accumulation in the brain such as  
CC Alzheimer's disease and Down's syndrome.  
CC See also AAR21932-75.  
XX  
SQ Sequence 11 AA;  
XX  
Query Match 95.1%; Score 58; DB 13; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 RPKPQOFFGLM 11  
DB 1 RPKPQqffgl 11  
XX  
RESULT 84  
ID AAR21951 standard; Peptide; 11 AA.  
XX  
AC AAR21951;  
XX  
DT 25-JUN-1992 (first entry)  
DE Substance P [Glu 3].  
XX  
KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX  
OS Synthetic.  
XX  
PN W09202248-A.  
XX  
PD 20-FEB-1992.  
XX  
PF 29-JUL-1991; 91WO-US05323.  
XX  
PR 27-JUL-1990; 90US-0559173.  
XX  
PA (CHTL-) CHILDRENS MED CENT.  
XX  
PI Yankner BA;  
XX  
DR WPI; 1992-079804/10.  
XX  
PT Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX  
PS Claim 10; Page 21; 35pp; English.  
XX  
CC The peptide is the tachykinin agonist substance P with a glutamic  
CC acid residue substituted at position 5. The peptide was  
CC synthesised by standard solid phase synthesis. Neuronal  
CC accumulation of beta-amyloid may be treated by administration of  
CC tachykinin agonists. The peptide can reduce the neurotoxic effects  
CC of a beta-amyloid related polypeptide on cultured neurons. The  
CC peptide and its analogues are useful for controlling diseases  
CC characterised by beta amyloid accumulation in the brain such as

CC Alzheimer's disease and Down's syndrome.  
CC See also AAR21932-75.  
XX  
SQ Sequence 11 AA;  
XX  
Query Match 95.1%; Score 58; DB 13; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 RPKPQOFFGLM 11  
DB 1 RPKPqffglm 11  
XX  
RESULT 85  
ID AAR28445 standard; peptide; 11 AA.  
XX  
AC AAR28445;  
XX  
DT 22-MAR-1993 (first entry)  
DE Neurokinine 1 ligand #3.  
XX  
KW NK1 receptor; tumour; malignant glioma; pheochromocytoma;  
KW paraganglia; small cell lung cancer; nerve regeneration; lymphoma;  
KW granuloma; Crohn's disease.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 11 /note= "amidated"  
FT  
XX  
PN W09218536-A.  
XX  
PD 29-OCT-1992.  
XX  
PF 22-APR-1992; 92WO-US03307.  
XX  
PR 22-APR-1991; 91EP-0200955.  
XX  
PA (MLCW ) MALLINCKRODT MEDICAL INC.  
XX  
PI Bakker WH, Hagen PM, Krenning EP, Lamberts SWJ, Visser TJ;  
XX  
DR WPI; 1992-382047/46.  
XX  
PT Detection and localisation of tissues with neurokinine-1 receptors  
PT - for detecting and treating tumours having neurokinine-1  
PT receptors e.g. malignant glioma, small cell lung cancer etc.  
XX  
PS Disclosure; Page 4; 22pp; English.  
XX  
CC This peptide or its Tyr0 deriv. is a preferred peptide having a  
CC selective affinity to neurokinine-1 receptors which (when  
CC labelled with a radioactive isotope) can be used in imaging methods.  
CC A generic formula for preferred peptides is AAR28441. Such peptides  
CC are thus useful in diagnosis and treatment of conditions that are  
CC related to NK1 receptors and in visualising NK1 receptors on certain  
CC tissues. See AAR28442-R28446.  
XX  
SQ Sequence 11 AA;  
XX  
Query Match 95.1%; Score 58; DB 13; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 RPKPQOFFGLM 11  
DB 1 RPKPqffglm 11

## RESULT 86

AA042649  
ID AAR42649 standard; peptide; 11 AA.

XX AC AAR42649;

XX DT 19-APR-1994 (first entry)

XX DE Neurokinin 1 receptor affinity-contg. peptide.

XX KW Neurokinin 1; somatostatin; receptor; cytokine; growth factor;  
KW hormone; intra-operativ; tumour; low energy gamma photon;  
KW radionuclide.

XX OS Synthetic.

XX FH Key Location/Qualifiers  
FT Modified-site 11

FT /note= "the C-terminal is amidated"

PN W09318797-A.

XX XX 30-SEP-1993.

XX XX 24-MAR-1993; 93WO-US02772.

XX XX 25-MAR-1992; 92EP-0200848.

XX PA (MLCW ) MALLINCKRODT MEDICAL INC.

XX PI Doedens BJ, Ensing GJ, Panek KJ;

XX XX WPI; 1993-320461/40.

XX XX Intra-operatively detecting and locating tumour tissues - using  
PT specific peptide(s) labelled with low energy gamma photon  
PT emitting radionuclide

XX PS Disclosure; Page 5; 31pp; English.

XX CC The method of intraoperatively detecting and locating tumoral  
CC tissues makes use of peptides having selective neurokinin 1  
CC receptor affinity (AAR42644; generic formula; AAR42646-R42650;  
CC specific examples), peptides having selective somatostatin  
CC receptor affinity (AAR42645; generic formula; AAR42651-R42660;  
CC specific examples), and peptides selected from cytokines,  
CC growth factors and hormones.

XX SQ Sequence 11 AA;

Query Match 95.1%; Score 58; DB 14; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11

Db 1 rpkpqfygim 11

## RESULT 87

AAW09003

ID AAW09003 standard; peptide; 11 AA.

XX AC AAW09003;

XX DT 03-MAR-1997 (first entry)

XX DE Substance P analogue, acts as substance P antagonist.

XX XX Analogue; substance P; spantide; non-peptide bond;

KW competitive inhibitor; receptor; neurogenic inflammation;  
KW rheumatoid arthritis; ulcerative colitis; eczema; Crohn's disease;  
KW anti-proliferative agent; small cell lung carcinoma; fibroblast.  
XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 6..7

FT /label= Gln-psi(CH2-NH)-Phe  
FT /note= "Opt. non-peptide linkage"

FT Modified-site 7..8

FT /label= Phe-psi(CH2-NH)-Phe  
FT /note= "Opt. non-peptide linkage"

FT Modified-site 8..9

FT /label= Phe-psi(CH2-NH)-Gly  
FT /note= "Opt. non-peptide linkage"

FT Modified-site 9..10

FT /label= Gly-psi(CH2-NH)-Leu  
FT /note= "Opt. non-peptide linkage"

FT Modified-site 10..11

FT /label= Leu-psi(CH2-NH)-Leu  
FT /note= "Position of claimed non-peptide linkage"

FT Modified-site 11

FT /note= "Amidated C-terminal"

XX XX US5410019-A.

XX XX 25-APR-1995.

XX PF 24-SEP-1987; 87US-0100571.

XX PR 30-MAR-1992; 92US-0860675.

XX PR 24-SEP-1987; 87US-0100571.

XX PR 25-MAR-1988; 88US-017311.

XX PR 08-JUN-1988; 88US-0204171.

XX PR 16-JUN-1988; 88US-0207759.

XX PR 23-SEP-1988; 88US-0248771.

XX PR 14-OCT-1988; 88US-0257998.

XX PR 09-DEC-1988; 88US-0282328.

XX PR 02-MAR-1989; 89US-0317941.

XX PR 16-AUG-1989; 89US-0394727.

XX XX (TULA ) TULANE EDUCATIONAL FUND.

XX PI Coy DH, Moreau J;

XX DR WPI; 1995-169633/22.

XX PT Novel linear peptide substance P analogues - useful as substance P  
PT antagonists, for treating neurogenic inflammation

XX PS Claim 3; Column 19; 16pp; English.

XX CC The sequences given in AAW09003-04 represent analogues of substance P  
CC and spantide, respectively. These analogues comprise a non-peptide  
CC bond between an amino acid residue of the active site, which occurs  
CC in the C-terminal half of the peptide, and an adjacent amino acid  
CC residue. They act as competitive inhibitors of the naturally  
CC occurring peptide by binding to its receptor. These peptides may be  
CC used in the treatment of diseases involving neurogenic inflammation,  
CC e.g. rheumatoid arthritis, ulcerative colitis, eczema and Crohn's  
CC disease. They are also useful as anti-proliferative agents, in  
CC the treatment of small cell lung carcinoma or disorders involving the  
CC proliferation of fibroblasts.

XX SQ Sequence 11 AA;

Query Match 95.1%; Score 58; DB 16; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RPKPQQFFGLM 11

Db 1 rpkpqgffglm 11  
|||||

RESULT 88  
AAW33181  
ID AAW33181 standard; peptide; 11 AA.

XX AC AAW33181;

XX DT 29-JAN-1998 (first entry)

XX DE Mono-DTPA-Lys1 Substance P.

XX KW Substance P; radiolabel; diagnostic imaging; therapy;  
KW mono-DTPA-Lys1.

XX OS Synthetic.

XX FH Key Location/Qualifiers

FT Modified-site 1

FT Modified-site /note= "DTPA-Lys"

FT Modified-site 11

FT Modified-site /note= "amidated"

XX PN W09640292-A1.

XX PD 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US09706.

XX PR 07-JUN-1995; 95US-0480372.

XX PA (MLCW ) MALLINCKRODT MEDICAL INC.

XX PI Srinivasan A;

XX DR WPI; 1997-087027/08.

XX PT Prepn. of pure radio-labelled peptide, e.g. for diagnostic imaging -  
PT by combining protected poly(amino:carboxylate) ligand with peptide  
PT and forming complex with radionuclide

XX PS Example 4; Page 12; 20pp; English.

XX CC Preparing a radiolabelled peptide composition, comprises combining  
CC a triamine or diamine chelating agent with a peptide, e.g. the  
CC present peptide, in a solid phase peptide synthesiser, and  
CC complexing a radionuclide with the chelate-peptide conjugate.  
CC Radiolabelled peptides or peptidomimetics can be used as diagnostic  
CC imaging agents, or in therapeutic applications, e.g. iodine(111)  
CC labelled pentareotide can be used for somatostatin receptor  
CC imaging of neuroendocrine tumours. The radiolabelled products are  
CC obtained efficiently and inexpensively in high purity. The  
CC protected polyaminocarboxylate ligands can be added to the peptide  
CC by standard solution or solid phase peptide synthesis and  
CC deprotected with conventional reagents to give only the  
CC mono-addition product, free of di-addition product impurities. The  
CC deprotected product can be labelled with medically useful  
CC radionuclides, e.g. lanthanides or actinides, at any desired  
CC location. Pre-derivatisation of individual amino acids is not  
CC required.

XX Sequence 11 AA;

Query Match 95.1%; Score 58; DB 18; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.00095;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQOFFGLM 11

Db 1 kpkpqgffglm 11

RESULT 89

AAW79775

ID AAW79775 standard; peptide; 11 AA.

XX AC AAW79775;

XX DT 07-JAN-1999 (first entry)

XX DE Substance P.

XX KW Tachykinin; neurokinin; NK1; receptor; antagonist; cystic fibrosis;

XX KW Substance P.

XX OS Mammalia.

XX PN US5830854-A.

XX PD 03-NOV-1998.

XX PF 14-DEC-1993; 93US-0166437.

XX PR 14-DEC-1992; 92GB-0026056.

XX PR 14-DEC-1992; 92GB-0026047.

XX PA (MERI ) MERCK SHARP & DOHME LTD.

XX PI Hargreaves RJ;

XX DR WPI; 1998-609287/51.

XX PT Treatment of cystic fibrosis - comprises administration of  
PT tachykinin receptor antagonist which is a neurokinin-1 receptor  
PT antagonist

XX PS Disclosure; Column 1; 12pp; English.

XX CC The invention relates to the new use of tachykinin receptor antagonists  
CC (particularly NK1 receptor antagonists) for the treatment of cystic  
CC fibrosis. The present sequence is that of Substance P, one of three  
CC known mammalian tachykinins.

XX Sequence 11 AA;

Query Match 95.1%; Score 58; DB 19; Length 11;

Best Local Similarity 90.9%; Pred. No. 0.00095;

Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQOFFGLM 11

Db 1 rpkpqgffglm 11

RESULT 90

AAW99689

ID AAW99689 standard; peptide; 11 AA.

XX AC AAW99689;

XX DT 03-JUN-1999 (first entry)

XX DE Substance P analogue #6.

XX KW Substance P receptor antagonist; analgesic; inhibitor; NMDA blocker;  
XX KW non-toxic N-methyl-D-aspartate receptor antagonist; muscular pain;  
XX KW musculoskeletal pain; chronic pain; neuropathic pain; migraine.

XX OS Synthetic.

XX FH Key

FT Modified-site

FT Location/Qualifiers

FT 10..11

```
FT Modified-site /note= "Leu-psi(CH2-NH)-Leu"  
FT 11  
FT /note= "amidated"  
XX WO9907413-A1.  
XX  
XX 18-FEB-1999.  
XX  
XX 26-MAY-1998; 98WO-US10707.  
XX  
XX 11-AUG-1997; 97US-0055233.  
XX  
XX (ALGO-) ALGOS PHARM CORP.  
XX  
XX Caruso FS;  
XX  
XX WPI; 1999-167216/14.  
XX  
XX New analgesic composition comprises - a substance P receptor  
PT antagonist with a substance P receptor antagonist potentiator, used  
PT for the treatment of pain  
XX  
XX Claim 3; Page 29; 54pp; English.  
XX  
XX A method has been developed for treating pain with: (a) a substance P  
CC receptor antagonist; and (b) a substance P receptor antagonist  
CC potentiator, i.e. N-methyl-D-aspartate (NMDA) receptor antagonist or  
CC substance that blocks at least 1 major intracellular consequence of  
CC NMDA receptor activation. The method can be used for treating muscular,  
CC musculoskeletal, chronic or neuropathic pain, or migraine. The present  
CC sequence represents a substance P analogue for use in the method.  
XX  
XX Sequence 11 AA;  
SQ  
  
Query Match 95.1%; Score 58; DB 20; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RPKQQQFFGLM 11  
DB |||||  
1 rpkqqffgll 11  
  
RESULT 91  
AAW92679  
ID AAW92679 standard; peptide; 11 AA.  
XX  
AC AAW92679;  
XX  
DT 30-APR-1999 (first entry)  
XX  
DE Human tachykinin agonist beta-amyloid peptide fragment #25.  
XX  
KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX  
OS Homo sapiens.  
XX  
PN US5876948-A.  
XX  
PD 02-MAR-1999.  
XX  
DT 30-APR-1999 (first entry)  
XX  
DE Human tachykinin agonist beta-amyloid peptide fragment #25.  
XX  
KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX  
OS Homo sapiens.  
XX  
PN US5876948-A.  
XX  
PD 02-MAR-1999.  
XX  
PF 27-JUL-1991; 91US-0737371.  
XX  
XX 29-JUL-1991; 91US-0737371.  
PR 27-JUL-1990; 90US-0559173.  
XX  
XX (CHIL-) CHILDRENS MEDICAL CENT.  
PA  
XX Yankner BA;  
PI  
XX  
XX WPI; 1999-189630/16.  
XX  
XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX  
XX Disclosure; Column 15-16; 28pp; English.  
XX  
XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage,  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
CC beta-amyloid peptide fragments.  
XX  
XX Sequence 11 AA;  
SQ  
  
Query Match 95.1%; Score 58; DB 20; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RPKQQQFFGLM 11  
DB |||||  
1 rpkqqffgll 11  
  
RESULT 92  
AAW92666  
ID AAW92666 standard; peptide; 11 AA.  
XX  
AC AAW92666;  
XX  
DT 30-APR-1999 (first entry)  
XX  
DE Human tachykinin agonist beta-amyloid peptide fragment #12.  
XX  
KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX  
OS Homo sapiens.  
XX  
PN US5876948-A.  
XX  
PD 02-MAR-1999.  
XX  
DT 27-JUL-1991; 91US-0737371.  
XX  
XX 29-JUL-1991; 91US-0737371.  
PR 27-JUL-1990; 90US-0559173.  
XX  
XX (CHIL-) CHILDRENS MEDICAL CENT.  
PA  
XX Yankner BA;  
PI  
XX  
XX WPI; 1999-189630/16.  
XX  
XX Screening for neurotoxin inhibitors - by testing compounds for their  
PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX  
XX Disclosure; Column 15-16; 28pp; English.  
XX  
XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage,  
CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human
```

CC beta-amyloid peptide fragments.

XX Sequence 11 AA;

Query Match 95.1%; Score 58; DB 20; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
|||:|||||  
Db 1 rpkpqyfglm 11

## RESULT 93

AAB91402  
ID AAB91402 standard; Peptide; 11 AA.

XX AC AAB91402;

XX DT 22-JUN-2001 (first entry)

XX DE Tachykinins peptide SEQ ID NO:578.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
XX KW blood component; modification; succinimidyl; maleimido group; amino;  
XX KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200069900-A2.

XX PD 23-NOV-2000.

XX PF 17-MAY-2000; 2000WO-US13576.

XX PR 17-MAY-1999; 99US-0134406.

XX PR 10-SEP-1999; 99US-0153406.

XX PR 15-OCT-1999; 99US-0159783.

XX PA (CONJ-) CONJUCHEM INC.

XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX DR WPI; 2001-112059/12.

XX PT Modifying and attaching therapeutic peptides to albumin prevents  
XX PT peptidase degradation, useful for increasing length of in vivo activity

XX PS Disclosure; Page 389; 733pp; English.

XX CC The present invention describes a modified therapeutic peptide (I)  
XX CC comprising a therapeutically active amino acid region (III) and a  
XX CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
XX CC a less therapeutically active amino acid region (IV), which covalently  
XX CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX CC factors and neurotransmitters, to protect them from peptidase activity  
XX CC in vivo for the treatment of various disorders. Endogenous therapeutic  
XX CC peptides are not suitable as drug candidates as they require frequent  
XX CC administration due to rapid degradation by peptidases in the body.  
XX CC Modifying and attaching therapeutic peptides to albumin prevents or  
XX CC reduces the action of peptidases to increase length of activity (half  
XX CC life) and specificity as bonding to large molecules decreases  
XX CC intracellular uptake and interference with physiological processes.  
XX CC AAB90829 to AAB92441 represent peptides which can be used in the  
XX CC exemplification of the present invention.

XX Sequence 11 AA;

Query Match 95.1%; Score 58; DB 22; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Query Match 95.1%; Score 58; DB 22; Length 11;  
Best Local Similarity 90.9%; Pred. No. 0.00095;  
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGLM 11  
|||:|||||  
Db 1 rrpqgffglm 11

## RESULT 94

AAB91409  
ID AAB91409 standard; Peptide; 11 AA.

XX AC AAB91409;

XX DT 22-JUN-2001 (first entry)

XX DE Tachykinins peptide SEQ ID NO:585.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
XX KW blood component; modification; succinimidyl; maleimido group; amino;  
XX KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200069900-A2.

XX PD 23-NOV-2000.

XX PF 17-MAY-2000; 2000WO-US13576.

XX PR 17-MAY-1999; 99US-0134406.

XX PR 10-SEP-1999; 99US-0153406.

XX PR 15-OCT-1999; 99US-0159783.

XX PA (CONJ-) CONJUCHEM INC.

XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX DR WPI; 2001-112059/12.

XX PT Modifying and attaching therapeutic peptides to albumin prevents  
XX PT peptidase degradation, useful for increasing length of in vivo activity

XX PS Disclosure; Page 391; 733pp; English.

XX CC The present invention describes a modified therapeutic peptide (I)  
XX CC comprising a therapeutically active amino acid region (III) and a  
XX CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
XX CC a less therapeutically active amino acid region (IV), which covalently  
XX CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
XX CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
XX CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
XX CC factors and neurotransmitters, to protect them from peptidase activity  
XX CC in vivo for the treatment of various disorders. Endogenous therapeutic  
XX CC peptides are not suitable as drug candidates as they require frequent  
XX CC administration due to rapid degradation by peptidases in the body.  
XX CC Modifying and attaching therapeutic peptides to albumin prevents or  
XX CC reduces the action of peptidases to increase length of activity (half  
XX CC life) and specificity as bonding to large molecules decreases  
XX CC intracellular uptake and interference with physiological processes.  
XX CC AAB90829 to AAB92441 represent peptides which can be used in the  
XX CC exemplification of the present invention.

XX Sequence 11 AA;

QY 1 RPKPQQFFGLM 11  
 |||||:||||  
 Db 1 rpkpqffgglm 11

## RESULT 95

AAAP50633  
 ID AAP50633 standard; Peptide; 10 AA.

AC AAP50633;

XX 09-MAR-1992 (first entry)

XX Substance P-like peptide, P2-11.

XX Hair tonic; growth; regeneration.

XX Synthetic.

XX JP60202807-A.

XX 14-OCT-1985.

XX 28-MAR-1984; 84JP-0058390.

XX 28-MAR-1984; 84JP-0058390.

XX (WEIJ ) MEIJI SEIKA KAISHA.

XX WPI; 1985-293619/47.

XX Hair tonic compsn. - comprises peptide contg. pyroglutamic acid

PT or other aminoacid(s) residue

XX Disclosure; Page 2; 3pp; Japanese.

XX The C-terminal residues 1-4 may be absent (P6-11, P5-11, P4-11 and  
 CC P3-11 respectively). The C-terminal is amidated. Substance P  
 CC (H-RKPEEFGLM-NH2) or these peptides derived from it can be used in  
 CC aq. soln. or suspension to promote hair growth and regeneration.  
 CC See also AAP50632 and AAP50634.

XX Sequence 10 AA;

Query Match 91.8%; Score 56; DB 6; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0019;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PKPQQFFGLM 11  
 |||||:||||  
 Db 1 pkpqffgglm 10

## RESULT 96

AAR21933

ID AAR21933 standard; Protein; 10 AA.

XX AAR21933;

XX 25-JUN-1992 (first entry)

XX Substance P (2-11) fragment.

XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
 XX syndrome; hereditary cerebral haemorrhage.

XX Synthetic.

XX WO9202248-A.

XX 20-FEB-1992.

XX 29-JUL-1991; 91WO-US05323.  
 XX  
 PR 27-JUL-1990; 90US-0559173.  
 XX

XX (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using  
 PT tachykinin agonists e.g. substance P, physalamin and neurokinin  
 PT B, for treating Alzheimer's disease, Downs syndrome, etc.

XX Claim 9; Page 21; 35pp; English.

XX The peptide is a tachykinin agonist consisting of residues 2-11 of  
 CC substance P. The peptide was synthesised by standard solid phase  
 CC synthesis. Analogues of the peptide, with N-terminal deletions down  
 CC to substance P (7-11) were also synthesised. Neuronal accumulation of  
 CC beta-amyloid may be treated by administration of these tachykinin  
 CC agonists. The peptides reduce the neurotoxic effects of a beta-  
 CC amyloid related polypeptide on cultured neurons. The peptide and  
 CC its analogues are useful for controlling diseases characterised by  
 CC beta amyloid accumulation in the brain such as Alzheimer's disease  
 CC and Down's syndrome.  
 CC See also AAR21932-75.

XX Sequence 10 AA;

Query Match 91.8%; Score 56; DB 13; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0019;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PKPQQFFGLM 11  
 |||||:||||  
 Db 1 pkpqffgglm 10

## RESULT 97

AAW92663

ID AAW92663 standard; peptide; 10 AA.

XX AAW92663;

XX 30-APR-1999 (first entry)

XX Human tachykinin agonist beta-amyloid peptide fragment #9.

XX Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
 XX Alzheimer's disease; Down's syndrome; amyloidosis; human;  
 XX hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.

XX Homo sapiens.

XX US5876948-A.

XX 02-MAR-1999.

XX 27-JUL-1991; 91US-0737371.

XX 29-JUL-1991; 91US-0737371.

XX 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MEDICAL CENT.

XX Yankner BA;

XX WPI; 1999-189630/16.

XX Screening for neurotoxin inhibitors - by testing compounds for their



PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells

PS Disclosure; Column 13-14; 28pp; English.

XX This invention describes a method for screening compounds for inhibiting  
CC a neurotoxin. The method involves incubating tachykinin agonists with  
CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
CC used for identifying compounds for treating diseases characterised by an  
CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
CC with amyloidosis, and non-inherited congenital angiodystrophy with cerebral  
CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
XX beta-amyloid peptide fragments.

SQ Sequence 10 AA;

Query Match 91.8%; Score 56; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PKPQOFFGLM 11  
Db 1 pkpqqffglm 10

RESULT 98

AAB91423  
ID AAB91423 standard; Peptide; 10 AA.

XX AC AAB91423;

XX DT 22-JUN-2001 (first entry)

XX DE Tachykinins peptide SEQ ID NO:599.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidy; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200069900-A2.

XX PD 23-NOV-2000.

XX PF 17-MAY-2000; 2000WO-US13576.

XX PR 17-MAY-1999; 99US-0134406.

XX PR 10-SEP-1999; 99US-0153406.

XX PR 15-OCT-1999; 99US-0159783.

XX PA (CONJ-) CONJUCHEM INC.

XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX DR WPI; 2001-112059/12.

PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity

PS Disclosure; Page 395; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidy and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptide stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity

CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.

SQ Sequence 10 AA;

Query Match 91.8%; Score 56; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQOFFGL 10  
Db 1 rpkgqffgl 10

RESULT 99

AAB91427

ID AAB91427 standard; Peptide; 10 AA.

XX AC AAB91427;

XX DT 22-JUN-2001 (first entry)

XX DE Tachykinins peptide SEQ ID NO:603.

XX KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidy; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX OS Homo sapiens.

XX OS Synthetic.

XX PN WO200069900-A2.

XX PD 23-NOV-2000.

XX PF 17-MAY-2000; 2000WO-US13576.

XX PR 17-MAY-1999; 99US-0134406.

XX PR 10-SEP-1999; 99US-0153406.

XX PR 15-OCT-1999; 99US-0159783.

XX PA (CONJ-) CONJUCHEM INC.

XX PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX DR WPI; 2001-112059/12.

PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity

PS Disclosure; Page 396; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidy and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptide stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or

CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.

XX Sequence 10 AA;

Query Match 91.8%; Score 56; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPKPQQFFGL 10  
|||||  
Db 1 rpqpqffgl 10

RESULT 100  
AAB91445  
ID AAB91445 standard; Peptide; 10 AA.

XX

AC AAB91445;

XX 22-JUN-2001 (first entry)

XX Tachykinins peptide SEQ ID NO:621.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;

XX blood component; modification; succinimidy; maleimido group; amino;

XX hydroxyl; thiol; hormone; growth factor; neurotransmitter.

XX Homo sapiens.

XX Synthetic.

XX WO200069900-A2.

XX 23-NOV-2000.

XX 17-MAY-2000; 2000WO-US13576.

XX 17-MAY-1999; 99US-0134406.

XX 10-SEP-1999; 99US-0153406.

XX 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;

XX WPI; 2001-112059/12.

XX Modifying and attaching therapeutic peptides to albumin prevents

XX peptidase degradation, useful for increasing length of in vivo activity

XX .

XX Disclosure; Page 402; 733pp; English.

XX The present invention describes a modified therapeutic peptide (I)

XX comprising a therapeutically active amino acid region (iii) and a

XX reactive group (ii) (e.g. succinimidy and maleimido groups) attached to

XX a less therapeutically active amino acid region (iv), which covalently

XX bonds with amino/hydroxyl/thiol groups on blood components to form a

XX peptidase stabilised therapeutic peptide composed of 3-50 amino acids.

XX (I) are useful for modifying therapeutic peptides e.g. hormones, growth

XX factors and neurotransmitters, to protect them from peptidase activity

XX in vivo for the treatment of various disorders. Endogenous therapeutic

XX peptides are not suitable as drug candidates as they require frequent

XX administration due to rapid degradation by peptidases in the body.

XX Modifying and attaching therapeutic peptides to albumin prevents or

XX reduces the action of peptidases to increase length of activity (half

XX life) and specificity as bonding to large molecules decreases

XX intracellular uptake and interference with physiological processes.

XX AAB90829 to AAB92441 represent peptides which can be used in the

CC exemplification of the present invention.

XX Sequence 10 AA;

Query Match 91.8%; Score 56; DB 22; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0019;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PKPQOFFGLM 11  
|||||  
Db 1 pkpqffglm 10

RESULT 101

AAR21945

ID AAR21945 standard; Protein; 11 AA.

XX

AC AAR21945;

XX 25-JUN-1992 (first entry)

XX Substance P [Pro 1].

XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;

XX syndrome; hereditary cerebral haemorrhage.

XX Synthetic.

XX WO9202248-A.

XX 20-FEB-1992.

XX 29-JUL-1991; 91WO-US05323.

XX 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using

XX tachykinin agonists e.g. substance P, physalaemin and neurokinin

XX B, for treating Alzheimer's disease, Downs syndrome, etc.

XX Claim 10; Page 21; 35pp; English.

XX The peptide is the tachykinin agonist substance P with a Proline

XX residue substituted at position 1. The peptide was

XX synthesised by standard solid phase synthesis. Neuronal

XX accumulation of beta-amyloid may be treated by administration of

XX tachykinin agonists. The peptide can reduce the neurotoxic effects

XX of a beta-amyloid related polypeptide on cultured neurons. The

XX peptide and its analogues are useful for controlling diseases

XX characterised by beta amyloid accumulation in the brain such as

XX Alzheimer's disease and Down's syndrome.

XX See also AAR21932-75.

XX Sequence 11 AA;

Query Match 91.8%; Score 56; DB 13; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 PKPQOFFGLM 11  
|||||  
Db 2 pkpqffglm 11

RESULT 102

AAR21936  
ID AAR21936 standard; Protein; 11 AA.  
XX  
AC AAR21936;

XX 25-JUN-1992 (first entry)

XX Substance P or (7-11) [Ethionine 11].

XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 11  
FT /label= OTHER  
FT /note= "Ethionine"

XX WO9202248-A.

XX 20-FEB-1992.

XX 29-JUL-1991; 91WO-US05323.

XX 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.

XX Claim 10; Page 21; 35pp; English.

XX The peptide is the tachykinin agonist substance P with an Ethionine  
CC residue substituted at position 11. The peptide was synthesised  
CC by standard solid phase synthesis. An N-terminal deleted peptide  
CC (7-11) with the same substitution was also synthesised. Neuronal  
CC accumulation of beta-amyloid may be treated by administration of  
CC tachykinin agonists. The peptides can reduce the neurotoxic effects  
CC of a beta-amyloid related polypeptide on cultured neurons. The  
CC peptide and its analogues are useful for controlling diseases  
CC characterised by beta amyloid accumulation in the brain such as  
CC Alzheimer's disease and Down's syndrome.  
CC See also AAR21932-75.

XX Sequence 11 AA;

Query Match 91.8%; Score 56; DB 13; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPKPQQFFGL 10

Db 1 rpkpqgffgl 10

RESULT 103

AAR21941  
ID AAR21941 standard; Protein; 11 AA.

XX AAR21941;

XX 25-JUN-1992 (first entry)

XX Substance P [pGLU 1].

XX

KW Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.  
XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1  
FT /label= OTHER  
FT /note= "OTHER = pyro Glu"

XX WO9202248-A.

XX 20-FEB-1992.

XX 29-JUL-1991; 91WO-US05323.

XX 27-JUL-1990; 90US-0559173.

XX (CHIL-) CHILDRENS MED CENT.

XX Yankner BA;

XX WPI; 1992-079804/10.

XX Treatment of neuronal accumulation of beta-amyloid - using  
PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
PT B, for treating Alzheimer's disease, Downs syndrome, etc.

XX Claim 10; Page 21; 35pp; English.

XX The peptide is the tachykinin agonist substance P with a pyro  
CC glutamic acid residue substituted at position 1. The peptide was  
CC synthesised by standard solid phase synthesis. Neuronal  
CC accumulation of beta-amyloid may be treated by administration of  
CC tachykinin agonists. The peptide can reduce the neurotoxic effects  
CC of a beta-amyloid related polypeptide on cultured neurons. The  
CC peptide and its analogues are useful for controlling diseases  
CC characterised by beta amyloid accumulation in the brain such as  
CC Alzheimer's disease and Down's syndrome.  
CC See also AAR21932-75.

XX Sequence 11 AA;

Query Match 91.8%; Score 56; DB 13; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PKPQQFFGLM 11

Db 2 pkpqgffglm 11

RESULT 104

AAR21944  
ID AAR21944 standard; Protein; 11 AA.

XX AAR21944;

XX 25-JUN-1992 (first entry)

XX Substance P [Pro 11].

XX Tachykinin agonist; beta-amyloid; Alzheimer's disease; Down's;  
KW syndrome; hereditary cerebral haemorrhage.

XX Synthetic.

XX WO9202248-A.

XX 20-FEB-1992.

XX 29-JUL-1991; 91WO-US05323.

XX PR 27-JUL-1990; 90US-0559173.  
XX XX (CHIL-) CHILDRENS MED CENT.  
XX PA Yankner BA;  
XX PI  
XX DR WPI; 1992-079804/10.  
XX XX  
XX PT Treatment of neuronal accumulation of beta-amyloid - using  
XX PT tachykinin agonists e.g. substance P, physalaemin and neurokinin  
XX PT B, for treating Alzheimer's disease, Downs syndrome, etc.  
XX XX  
XX PS Claim 10; Page 21; 35pp; English.  
XX CC The peptide is the tachykinin agonist substance P with a Proline  
XX CC residue substituted at position 11. The peptide was  
XX CC synthesised by standard solid phase synthesis. Neuronal  
XX CC accumulation of beta-amyloid may be treated by administration of  
XX CC tachykinin agonists. The peptide can reduce the neurotoxic effects  
XX CC of a beta-amyloid related polypeptide on cultured neurons. The  
XX CC peptide and its analogues are useful for controlling diseases  
XX CC characterised by beta amyloid accumulation in the brain such as  
XX CC Alzheimer's disease and Down's syndrome.  
XX CC See also AAR21932-75.  
XX XX  
XX SQ Sequence 11 AA;  
  
Query Match 91.8%; Score 56; DB 13; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RPKPQQFFGL 10  
DB 1 rpkpqffgl 10  
  
RESULT 105  
AAW92709  
ID AAW92709 standard; peptide; 11 AA.  
XX AC AAW92709;  
XX XX  
XX DT 30-APR-1999 (first entry)  
XX XX  
XX DE Human tachykinin agonist beta-amyloid peptide fragment #55.  
XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
XX KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
XX KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX XX  
XX OS Homo sapiens.  
XX XX  
XX FH Key Location/Qualifiers  
XX FT Modified-site 9  
XX FT /label= MeGly  
XX FT /note= "N-methyl-glycine (sarcosine)"  
XX FT Modified-site 11  
XX FT /note= "Residue is Met(O2)"  
XX XX  
XX PN US5876948-A.  
XX PD 02-MAR-1999.  
XX XX  
XX XX 27-JUL-1991; 91US-0737371.  
XX PF  
XX XX 29-JUL-1991; 91US-0737371.  
XX PR  
XX PR 27-JUL-1990; 90US-0559173.  
XX XX  
XX PA (CHIL-) CHILDRENS MEDICAL CENT.  
XX XX  
XX PI Yankner BA;

XX WPI; 1999-189630/16.  
XX XX  
XX PT Screening for neurotoxin inhibitors - by testing compounds for their  
XX PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX XX  
XX PS Disclosure; Column 35-36; 28pp; English.  
XX XX  
XX CC This invention describes a method for screening compounds for inhibiting  
XX CC a neurotoxin. The method involves incubating tachykinin agonists with  
XX CC neuronal cells and a beta-amyloid peptide neurotoxin. The methods can be  
XX CC used for identifying compounds for treating diseases characterised by an  
XX CC undesirable build up of beta-amyloid protein, e.g. Alzheimer's disease,  
XX CC Down's syndrome, and the syndromes of hereditary cerebral haemorrhage  
XX CC with amyloidosis and non-inherited congophilic angiopathy with cerebral  
XX CC haemorrhage. AAW92655-W92731 are tachykinin agonists derived from human  
XX CC beta-amyloid peptide fragments.  
XX XX  
XX SQ Sequence 11 AA;  
  
Query Match 91.8%; Score 56; DB 20; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 RPKPQQFFGL 10  
DB 1 rpkpqffgl 10  
  
RESULT 106  
AAW92717  
ID AAW92717 standard; peptide; 11 AA.  
XX AC AAW92717;  
XX XX  
XX DT 30-APR-1999 (first entry)  
XX XX  
XX DE Human tachykinin agonist beta-amyloid peptide fragment #63.  
XX KW Tachykinin agonist; beta-amyloid; inhibition; neurotoxin; treatment;  
XX KW Alzheimer's disease; Down's syndrome; amyloidosis; human;  
XX KW hereditary cerebral haemorrhage; non-inherited congophilic angiopathy.  
XX XX  
XX OS Homo sapiens.  
XX XX  
XX FH Key Location/Qualifiers  
XX FT Modified-site 11  
XX FT /label= MeMet  
XX FT /note= "N-methyl-methionine"  
XX XX  
XX PN US5876948-A.  
XX PD 02-MAR-1999.  
XX XX  
XX XX 27-JUL-1991; 91US-0737371.  
XX PF  
XX XX 29-JUL-1991; 91US-0737371.  
XX PR  
XX PR 27-JUL-1990; 90US-0559173.  
XX XX  
XX PA (CHIL-) CHILDRENS MEDICAL CENT.  
XX XX  
XX PI Yankner BA;  
XX XX  
XX DR WPI; 1999-189630/16.  
XX XX  
XX PT Screening for neurotoxin inhibitors - by testing compounds for their  
XX PT effect on beta-amyloid peptide neurotoxic effect on neuronal cells  
XX XX  
XX PS Disclosure; Column 37-38; 28pp; English.  
XX XX  
XX CC This invention describes a method for screening compounds for inhibiting  
XX CC a neurotoxin. The method involves incubating tachykinin agonists with